



TITLE:

# Intelligent Image-Activated Cell Sorting

AUTHOR(S):

Nitta, Nao; Sugimura, Takeaki; Isozaki, Akihiro; Mikami, Hideharu; Hiraki, Kei; Sakuma, Shinya; Iino, Takanori; ...  
Uemura, Sotaro; Ozeki, Yasuyuki; Goda, Keisuke

---

CITATION:

Nitta, Nao ...[et al]. Intelligent Image-Activated Cell Sorting. Cell 2018, 175(1): 266-276.e13

ISSUE DATE:

2018-09-20

URL:

<http://hdl.handle.net/2433/235563>

RIGHT:

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>; The full-text file will be made open to the public on 20 September 2019 in accordance with publisher's 'Terms and Conditions for Self-Archiving'. ; この論文は出版社版ではありません。引用の際には出版社版をご確認ください。 ; This is not the published version. Please cite only the published version.

Resource

## Intelligent Image-Activated Cell Sorting

Nao Nitta<sup>1,2</sup>

Takeaki Sugimura<sup>1,2</sup>

Akihiro Isozaki<sup>1</sup>

Hideharu Mikami<sup>1</sup>

Kei Hiraki<sup>1</sup>

Shinya Sakuma<sup>3</sup>

Takanori Iino<sup>4</sup>

Fumihito Arai<sup>3</sup>

Taichiro Endo<sup>5,6</sup>

Yasuhiro Fujiwaki<sup>4</sup>

Hideya Fukuzawa<sup>7</sup>

Misa Hase<sup>1</sup>

Takeshi Hayakawa<sup>8</sup>

Kotaro Hiramatsu<sup>1</sup>

Yu Hoshino<sup>9</sup>

Mary Inaba<sup>10</sup>

Takuro Ito<sup>1,2</sup>

Hiroshi Karakawa<sup>1</sup>

Yusuke Kasai<sup>3</sup>

Kenichi Koizumi<sup>10</sup>

SangWook Lee<sup>1</sup>

Cheng Lei<sup>1</sup>

Ming Li<sup>11</sup>

Takanori Maeno<sup>12</sup>

Satoshi Matsusaka<sup>13</sup>

Daichi Murakami<sup>10</sup>

Atsuhiko Nakagawa<sup>14</sup>

Yusuke Oguchi<sup>15</sup>

Minoru Oikawa<sup>16</sup>

Tadataka Ota<sup>1</sup>

Kiyotaka Shiba<sup>17</sup>

Hirofumi Shintaku<sup>18</sup>

Yoshitaka Shirasaki<sup>15</sup>

Kanako Suga<sup>17</sup>

Yuta Suzuki<sup>4</sup>

Nobutake Suzuki<sup>15</sup>

Yo Tanaka<sup>19</sup>

Hiroshi Tezuka<sup>10</sup>

Chihana Toyokawa<sup>7</sup>

Yaxiaer Yalikun<sup>19</sup>

Makoto Yamada<sup>5, 20</sup>

Mai Yamagishi<sup>15</sup>

Takashi Yamano<sup>7</sup>

Atsushi Yasumoto<sup>21</sup>

Yutaka Yatomi<sup>21</sup>

Masayuki Yazawa<sup>22</sup>

Dino Di Carlo<sup>1, 11, 23, 24</sup>

Yoichiroh Hosokawa<sup>25</sup>

Sotaro Uemura<sup>15</sup>

Yasuyuki Ozeki<sup>4</sup>

Keisuke Goda<sup>1, 2, 26, 27, \*</sup>

[goda@chem.s.u-tokyo.ac.jp](mailto:goda@chem.s.u-tokyo.ac.jp)

<sup>1</sup>Department of Chemistry, The University of Tokyo, Tokyo 113-0033, Japan

<sup>2</sup>Japan Science and Technology Agency, Saitama 332-0012, Japan

<sup>3</sup>Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Nagoya 464-8601, Japan

<sup>4</sup>Department of Electrical Engineering and Information Systems, The University of Tokyo, Tokyo 113-8656, Japan

<sup>5</sup>Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan

<sup>6</sup>ExaWizards Inc., Tokyo 105-0013, Japan

<sup>7</sup>Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan

<sup>8</sup>Department of Precision Mechanics, Chuo University, Tokyo 192-0393, Japan

<sup>9</sup>Department of Chemical Engineering, Kyushu University, Fukuoka 819-0395, Japan

- <sup>10</sup>Department of Creative Informatics, The University of Tokyo, Tokyo 113-0033, Japan
- <sup>11</sup>Department of Bioengineering, University of California, Los Angeles, Los Angeles, CA 90095, USA
- <sup>12</sup>Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan
- <sup>13</sup>Department of Gastroenterology, Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan
- <sup>14</sup>Department of Neurosurgery, Graduate School of Medicine, Tohoku University, Sendai 980-8577, Japan
- <sup>15</sup>Department of Biological Sciences, The University of Tokyo, Tokyo 113-0033, Japan
- <sup>16</sup>Science and Technology Unit, Natural Sciences Cluster, Kochi University, Kochi 780-8520, Japan
- <sup>17</sup>Division of Protein Engineering, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan
- <sup>18</sup>Department of Micro Engineering, Kyoto University, Kyoto 606-8501, Japan
- <sup>19</sup>Center for Biosystems Dynamics Research, RIKEN, Osaka 565-0871, Japan
- <sup>20</sup>Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
- <sup>21</sup>Department of Clinical Laboratory Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan
- <sup>22</sup>Department of Rehabilitation and Regenerative Medicine, Pharmacology, Columbia University, New York, NY 10032, USA
- <sup>23</sup>Department of Mechanical Engineering, University of California, Los Angeles, Los Angeles, CA 90095, USA
- <sup>24</sup>California NanoSystems Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA
- <sup>25</sup>Graduate School of Materials Science, Nara Institute of Science and Technology, Ikoma 630-0192, Japan
- <sup>26</sup>Department of Electrical Engineering, University of California, Los Angeles, Los Angeles, CA 90095, USA
- Corresponding author
- <sup>27</sup>Lead Contact

Published: August 27, 2018



## Summary

A fundamental challenge of biology is to understand the vast heterogeneity of cells, particularly how cellular composition, structure, and morphology are linked to cellular physiology. Unfortunately, conventional technologies are limited in uncovering these relations. We present a machine-intelligence technology based on a radically different architecture that realizes real-time image-based intelligent cell sorting at an unprecedented rate. This technology, which we refer to as intelligent image-activated cell sorting, integrates high-throughput cell microscopy, focusing, and sorting on a hybrid software-hardware data-management infrastructure, enabling real-time automated operation for data acquisition, data processing, decision-making, and actuation. We use it to demonstrate real-time sorting of microalgal and blood cells based on intracellular protein localization and cell-cell interaction from large heterogeneous populations for studying photosynthesis and atherothrombosis, respectively. The technology is highly versatile and expected to enable machine-based scientific discovery in biological, pharmaceutical, and medical sciences.

## Introduction

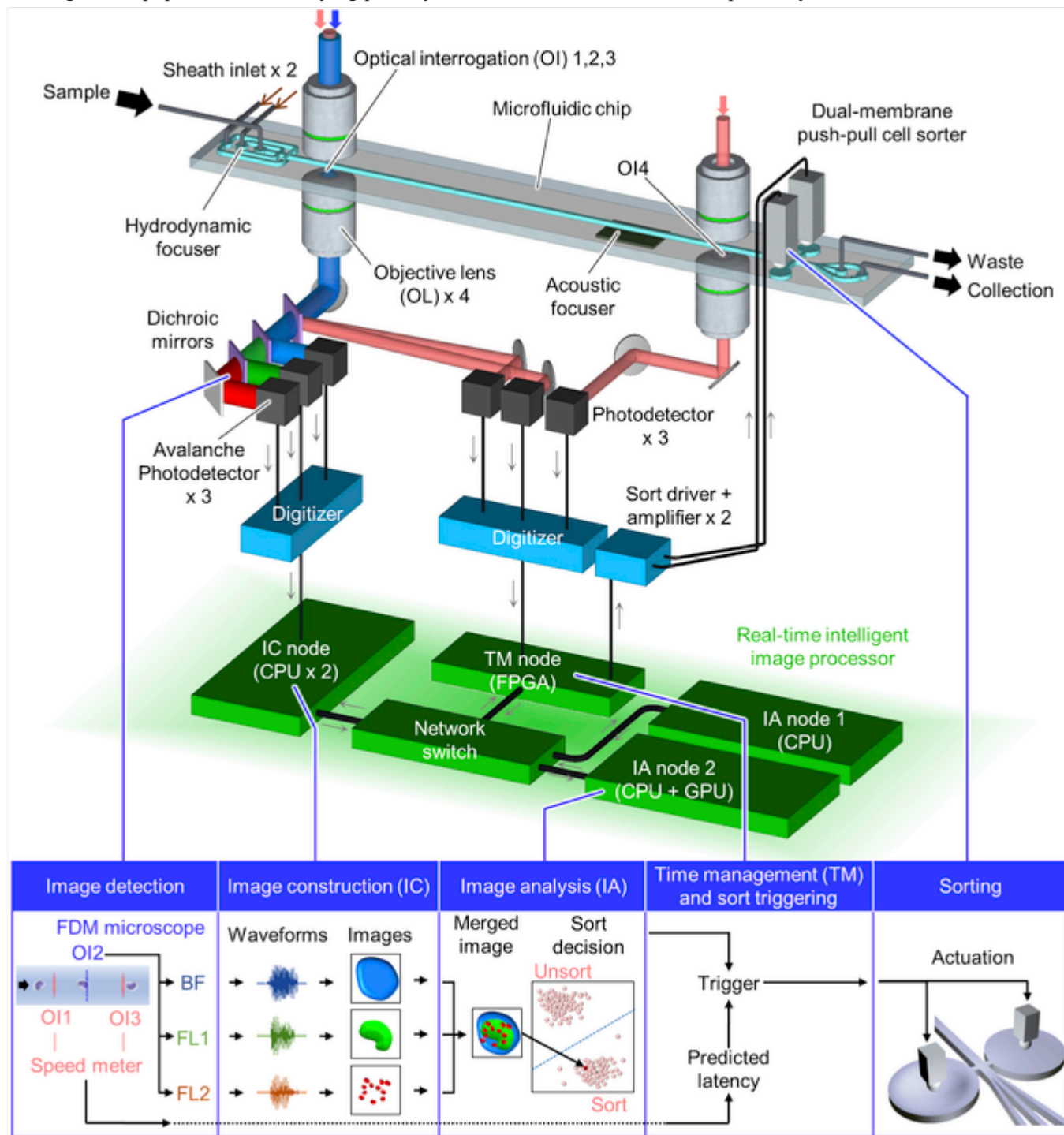
A fundamental challenge of biology is to study the vast differences that exist between cells, even those with identical genomes (Altschuler and Wu, 2010; Stubbington et al., 2017). The differences in composition, structure, and morphology of cells are connected to their physiological functions and are an important aspect of cell identity (Hao et al., 2013; von Erlach et al., 2018; Moor et al., 2017; Zenker et al., 2017; Boutros et al., 2015; Caicedo et al., 2017; Pernas et al., 2018; Mackinder et al., 2017). For example, localization of transcription factors to the nucleus or cytoplasm leads to dramatic differences in cellular behavior (Hao et al., 2013), while cellular geometry (e.g., size, shape) affects intracellular signaling and regulates cell growth and differentiation (von Erlach et al., 2018). A range of other morphological features such as nuclear shape, nucleus-to-cytoplasm ratio, cytoskeleton structure, chromosome abnormality, cell-cell interaction, RNA localization, and lipid droplet distribution are likely also linked to unique cellular function (Moor et al., 2017; Zenker et al., 2017; Boutros et al., 2015; Caicedo et al., 2017; Pernas et al., 2018; Mackinder et al., 2017). To understand such cell-to-cell differences, new approaches are needed to rapidly search through and sort out cells with unique chemical and morphological features from large heterogeneous populations (Altschuler and Wu, 2010; Chattopadhyay et al., 2014), as conventional technologies have been limited in uncovering these relations. We anticipate that a technology to sort cells based on their images at high rates will be foundational to uncovering new subpopulations and rare features of cells and accelerating progress toward an era of precision medicine

(Jiang et al., 2017), cell therapy (von Erlach et al., 2018), drug discovery (Boutros et al., 2015; Caicedo et al., 2017), metabolic engineering (Nielsen and Keasling, 2016; Wakisaka et al., 2016), and protein engineering (Piatkevich et al., 2018).

Unfortunately, realization of such a technology is significantly challenging due to the trade-off between the volume and complexity of image data (which is correlated to accuracy) and the data transfer and image-processing speed (which is correlated to response time). In fact, conventional high-throughput cell-sorting technologies such as fluorescence-activated cell sorting (FACS) (Tung et al., 2007) can only handle low-resolution data (e.g., ~20 light scattering and fluorescence signals without spatial information) for real-time data processing and actuation (i.e., sorting). On the other hand, digital analysis of high-resolution data (i.e., images) in image-based high-content screening (Boutros et al., 2015; Caicedo et al., 2017; Brasko et al., 2018) and imaging flow cytometry (George et al., 2006; Caicedo et al., 2017; Basiji et al., 2007; Lee et al., 2018), where machine learning is effective by virtue of the availability of the high-dimensional data (Eulenberg et al., 2017), is too slow to perform real-time actuation due to the limited speed of data transfer and image processing. A typical strategy for increasing the data-processing throughput is to process the data in parallel by using multiple computers, but this is generally limited to “offline” operation, resulting in a long turnaround time (more than several seconds), and does not allow real-time actuation. Previous efforts to achieve real-time analysis and sorting have employed unique electrical (Fabbri et al., 2013), chemical (Sun et al., 2014), microfluidic (Girault et al., 2017), optical (Ota et al., 2018), and mechanical (Brasko et al., 2018) approaches, but these approaches lack high-throughput or imaging capabilities. Presently, there is no high-throughput image-based cell-sorting technology that is validated and available.

In this Resource, we present a real-time machine-intelligence technology that overcomes the trade-off between accuracy and speed by employing a radically different data-management infrastructure and hence realizes real-time “online” image-based intelligent cell search and sorting at an unprecedented rate. The technology, which we refer to as the intelligent image-activated cell sorter (IACS), integrates high-throughput cell microscopy (Mikami et al., 2018), focusing, and sorting (Sakuma et al., 2017) on a hybrid software-hardware data-management infrastructure that runs a telecom-grade 10-Gbps all-Internet-protocol (IP) architecture (Figure 1) (Hiraki et al., 2016). This optical-microfluidic-electrical-computational-mechanical system enables high flexibility, high scalability, and automated operation for data acquisition, data processing, decision-making, and actuation, all within 32 ms even with deep-learning algorithms. In a sense, the intelligent IACS is an imaging version of FACS, which offers many more capabilities in both technological and applicational aspects. To show the power and broad utility of the intelligent IACS (e.g., for microbiology, hematology), we demonstrate real-time sorting of microalgal

and blood cells based on intracellular protein localization and cell-cell interaction from large heterogeneous populations for studying photosynthesis and atherothrombosis, respectively.



**Figure 1** Schematic and Functionality of the Intelligent IACS. Suspended cells in a sample injected into the intelligent IACS are focused by the hydrodynamic focuser into a single stream, imaged by the FDM microscope, analyzed by the real-time intelligent image processor, maintained in a single stream by the acoustic focuser during the computation, and sorted by the dual-membrane push-pull cell sorter triggered by decisions made by the image processor. The entire process is operated in a fully automated and real-time manner. Video S1 shows the full functionality of the intelligent IACS. OI1-OI4, optical interrogation points for the FDM microscope and speed meter; BF, bright field; FL1, FL2, fluorescence in different color bands. Also, see Figures S1A–S1D, S2, S3A–S3H, and S4A–S4F.

## Results

### Schematic and Functionality

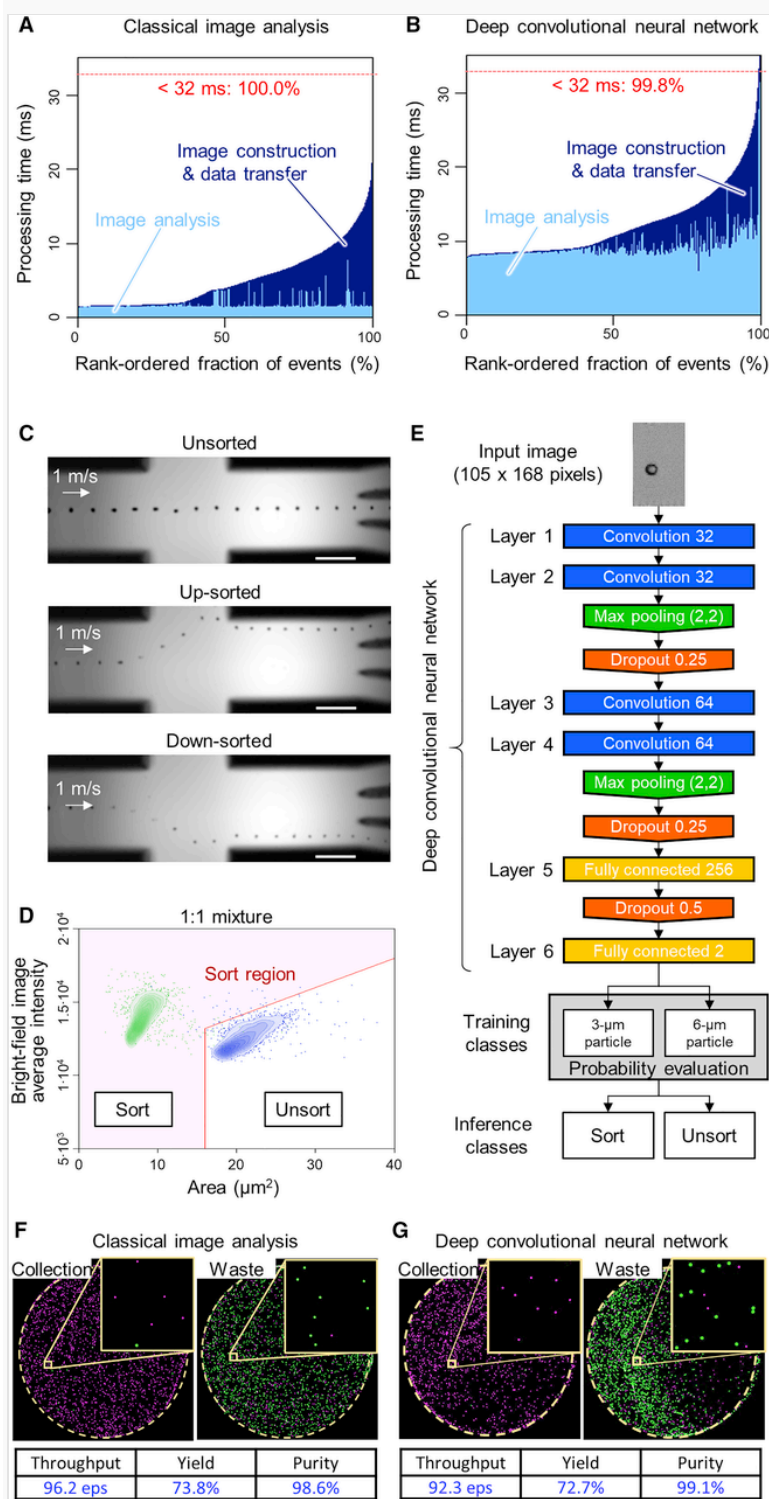
The intelligent IACS is schematically shown in Figure 1 (Figures S1A–S1D and S2; STAR Methods). It is a seamless integration of (1) a two-step three-dimensional (3D) on-chip hydrodynamic cell focuser (Figures S1A and S3A–S3D); (2) a frequency-division-multiplexed (FDM) microscope (Mikami et al., 2018) for continuous, high-speed, blur-free, sensitive, bright-field, and two-color fluorescence image acquisition of cells flowing at 1 m/s (Figures S1C, S1D, and S3E–S3H), which is required to achieve a high throughput of ~100 events per second (eps) as described below; (3) a speed meter for measuring the flow speed of cells and predicting their arrival time at the sort point (Figures S1D and S4A–S4D); (4) a real-time intelligent image processor composed of a field-programmable gate array (FPGA), three central processing units (CPUs), a graphics processing unit (GPU), and a network switch, all on a 10-Gbps all-IP network for high-speed digital image processing and decision-making (Figure S2); (5) a 3D on-chip acoustic cell focuser for maintaining a focused cell stream for a long distance (>3 cm) (Figure S1A; STAR Methods); and (6) an on-chip dual-membrane push-pull cell sorter for rapidly isolating target cells from the cell stream (Sakuma et al., 2017) (Figures S1A, S1B, and S1D). As shown in Video S1, suspended cells in a sample injected into the intelligent IACS are focused by the hydrodynamic focuser into a single stream, imaged by the FDM microscope, analyzed by the real-time intelligent image processor, maintained in the single stream by the acoustic focuser during the computation, and sorted by the dual-membrane push-pull cell sorter triggered by decisions made by the image processor. The entire process is operated in a fully automated and real-time manner.

The functionality of each critical element (1)–(6) is described as follows. First, the two-step 3D on-chip hydrodynamic focuser uses a sheath flow to confine flowing cells in the horizontal direction, followed by another sheath flow to confine them in the vertical direction, in order to produce a highly stable 3D focused cell flow at 1 m/s for blur-free image acquisition with the FDM microscope. A flow controller that consists of pressure vessels, electro-pneumatic regulators, and liquid flow meters is used to provide the optimum flow rate ratio between the sample, horizontal sheath, and vertical sheath flows (Figures S3A and S3B). Second, the FDM microscope images every single cell in the high-speed flow by employing the FDM image acquisition technique (Mikami et al., 2018). Specifically, it employs a linear array of multiple intensity-modulated excitation beams with distinct modulation frequencies. The beam array is focused across the microchannel to produce a linear array of focal spots in the direction perpendicular to the flow. Each flowing cell is scanned in the flow direction by the beam array to acquire a sequence of one-dimensional (1D) bright-field and fluorescence spatial profiles of the cell with single-pixel photodetectors. Two-dimensional (2D) bright-field and fluorescence images of the cell are constructed by stacking the 1D spatial profiles after demodulating the detected transmission and fluorescence signals at the modulation frequencies (Figures S3E–S3H). Third, the speed meter measures forward-scattered light from each flowing cell at the FDM microscope to predict the optimum sort timing (i.e., the sort latency at the sort point). The speed measurement is important in the intelligent IACS (not commonly conducted in FACS) because while in FACS it typically takes only a few hundred microseconds for a cell to flow from the optical interrogation point to the sort point, the sort latency of the intelligent IACS needs to be >100 times longer than that of FACS due to its requirement for analyzing the much larger amount of data (images) on the real-time intelligent image processor. Precise prediction of the sort latency is, hence, required for high-purity sorting (Figures S4A–S4D). Fourth, during the computation of the real-time intelligent image processor, the 3D on-chip acoustic cell focuser keeps flowing cells at the center of the microchannel all the way from the FDM microscope to the sort point (~3.2 cm) by exciting both horizontal and vertical acoustic resonance modes of the square-shaped cross section of the microchannel with piezoelectric transducers. It serves to avoid unwanted position shifts of flowing cells due to inertial lift and gravitational forces that would otherwise introduce cell-morphology-dependent fluctuations in the latency (Figures S3C and S3D). Fifth, the on-chip dual-membrane push-pull cell sorter rapidly controls local flow at the sort point to isolate target cells from the central stream with piezoelectrically actuated dual glass-membrane pumps (Sakuma et al., 2017). Deforming each glass membrane with the corresponding piezoelectric actuator produces high-speed local flow that crosses the central microchannel at the sort point. In the absence of the pump actuation, cells flow into the central

branch of the three-branch microchannel junction, whereas with the pump actuation at the right timing, cells flow into either the upper or lower branch.

The brain of the intelligent IACS (the final element that connects all the other elements) is the hybrid FPGA-CPU-GPU infrastructure on the 10-Gbps all-IP network (Hiraki et al., 2016) that enables high flexibility, high scalability, and real-time automated operation for intelligent image processing and decision-making (Figures 1 and S2; STAR Methods). An IP network is a communication network that uses the Internet protocol to send and receive signals between multiple computers, multiple sensors, and multiple actuators and is an essential element of the Internet of Things. This communication network acts as a platform for the FPGA, CPUs, and GPU to digitally process different types of data (three avalanche photodetector signals from the FDM microscope and three photodiode signals from the speed meter) and communicate among each other at a high speed of 10 Gbps. The CPUs and GPU perform image construction of one or more cells (events) based on the three-color image acquisition (represented by the IC node in Figure 1) and image analysis of the events (represented by the IA nodes in Figure 1), which includes multidimensional feature extraction, classification, and sort-decision-making, whereas the FPGA (represented by the TM node in Figure 1) precisely determines the optimum sort timing based on signals from the speed meter, predicts the latency at the sorter with a high precision of 200  $\mu$ s (Figure S4D; STAR Methods), and generates a trigger signal for the sort driver (another FPGA located outside the network) followed by amplifiers and piezoelectric transducers for the dual-membrane pumps if a target event is identified. In contrast to FACS, which only uses FPGA-based hardware for real-time signal processing, the combination of the FPGA-based hardware and the CPU-GPU-based software in the intelligent IACS offers much higher flexibility and scalability to scale and deploy on-demand network functions based on complex digital algorithms for image analysis, machine learning, and cell sorting as described in our experimental demonstrations below. Selection of digital algorithms is highly flexible within the range of the specified processing time, thereby rendering previously developed image-processing algorithms (including machine learning) for microscopy and imaging flow cytometry platforms (Caicedo et al., 2017; Basiji et al., 2007; Lee et al., 2018; Schindelin et al., 2012; Grys et al., 2017) readily available to the intelligent IACS. For the real-time automated operation, all the above functions (i.e., image construction, image analysis, data transfer) are conducted only within a total of 32 ms for both a classical algorithm (Figure 2A) and a deep convolutional neural network (CNN) on which deep learning is implemented (Figure 2B). Moreover, the sequence of these functions is processed in parallel for consecutive events by pipeline processing (i.e., pipelined like a bucket brigade), allowing for the unprecedented throughput of  $\sim$ 100 eps (Figures 1, S2, S4E, and S4F).





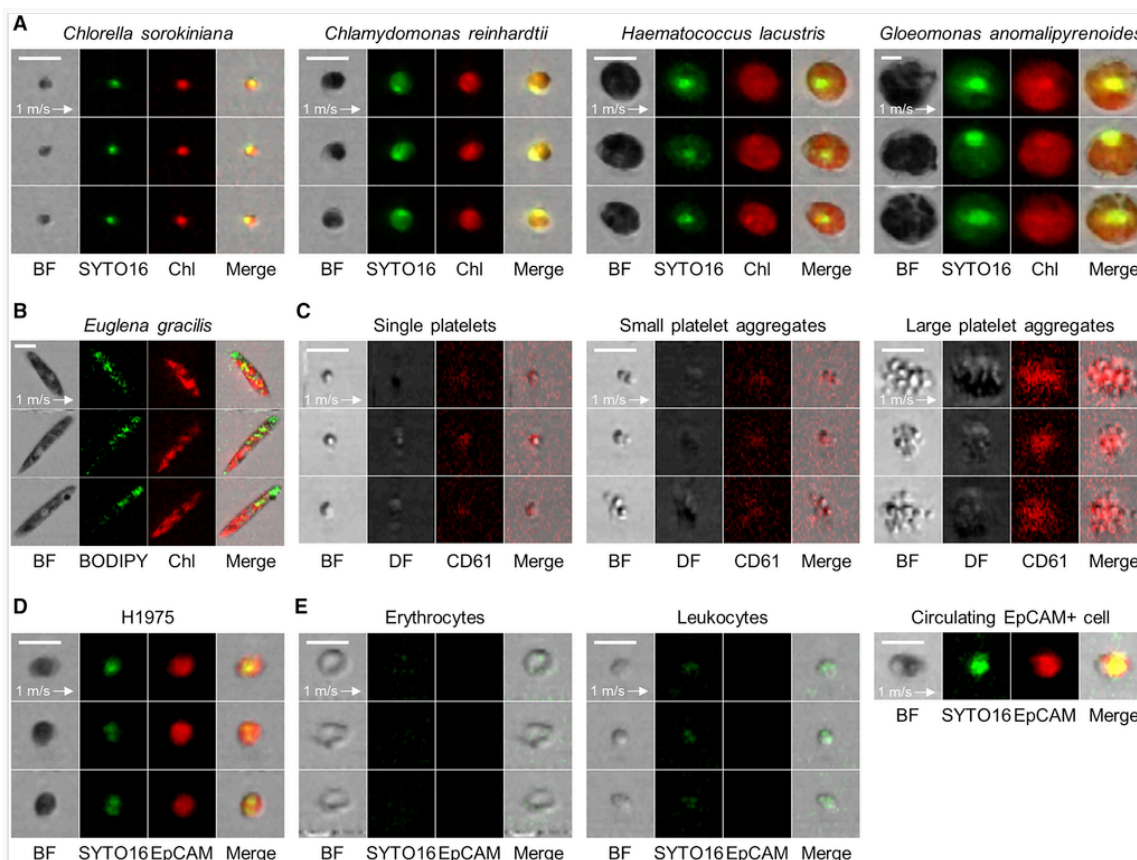
**Figure 2** Basic Performance of the Intelligent IACS(A and B) Processing time of the image construction and data transfer combined (dark blue) and the image analysis (light blue) for each event using 6- $\mu\text{m}$  particles, based on the classical algorithm and the deep CNN on which deep learning is implemented. The measured events are rank-ordered in the total processing time.(C) Trajectories of a 6- $\mu\text{m}$  particle (unsorted) and 3- $\mu\text{m}$  particles (sorted) flowing at 1 m/s, recorded by a high-speed CMOS camera at 20,000 frames/s.(D) Sort region to sort 3- $\mu\text{m}$  particles with the classical algorithm.(E) Deep CNN in a six-layer structure with four convolution layers and two fully connected layers for two training classes: 3- $\mu\text{m}$  particles and 6- $\mu\text{m}$  particles.(F and G) Enrichment of 3- $\mu\text{m}$  particles from a 1:1 mixture of 3- $\mu\text{m}$  and 6- $\mu\text{m}$  particles with the classical algorithm and the deep CNN. 3- $\mu\text{m}$  and 6- $\mu\text{m}$  particles are shown in purple and green, respectively. In both demonstrations, a high throughput of  $\sim 100$  eps and a high purity of  $\sim 99\%$  were achieved.

## Basic Performance

The basic performance of the intelligent IACS was validated as follows. First, we evaluated high-purity enrichment of fluorescent particles with a classical image analysis algorithm and a six-layer CNN on one of the IA nodes (Figures 1 and S2). Figure 2C shows the typical particle-sorting process of the intelligent IACS recorded by a high-speed CMOS image sensor. A 1:1 mixture of 3- $\mu\text{m}$  and 6- $\mu\text{m}$  particles was tested as a sample to demonstrate sorting of 3- $\mu\text{m}$  particles based on the classical image analysis algorithm (Figure 2D) and the CNN (Figure 2E). An average throughput of  $\sim 100$  eps was achieved continuously, while sorted and unsorted particles were simultaneously collected. Our microscopic enumeration of the sorted and unsorted particles clearly indicates the enrichment of the smaller particles with a high purity of 98.6% and 99.1% based on the classical algorithm (Figure 2F) and the CNN (Figure 2G), respectively. In addition, to evaluate the performance of the intelligent IACS in a biologically relevant situation, we tested a 1:19 mixture of 3- $\mu\text{m}$  and 6- $\mu\text{m}$  particles and sorted 3- $\mu\text{m}$  particles from the mixture with the classical image analysis algorithm with a high throughput, yield, and purity of 106.4 eps, 67.7%, and 96.2%, respectively. In these experiments, we optimized the settings of the intelligent IACS to obtain high purity in favor of yield. However, the yield can also be improved in favor of the purity or throughput. Second, to show the broad utility of the intelligent IACS from microbiology to hematology, we evaluated its imaging capability with diverse cell types. Figures 3A–3E show bright-field and two-color fluorescence images of various microalgal and human cells we tested. The intelligent IACS is also able to directly detect a highly rare EpCAM-positive circulating cell in blood from a patient with



pancreatic cancer (Figure 3E). The high quality of the images indicates that the intelligent IACS is capable of identifying the intracellular molecular distribution and morphological features of various cell types ranging from 3  $\mu\text{m}$  to 30  $\mu\text{m}$  in cell diameter. Among these cell types, below we used a model photosynthetic eukaryote, *Chlamydomonas reinhardtii*, and human platelets to show practical applications of the intelligent IACS, which are impossible to perform by FACS or would be extremely labor intensive or time consuming if they were conducted by microscopy with manual pipetting or even with a robotic arm.



**Figure 3** Various Types of Cells Imaged by the Intelligent IACS. The flow speed of all the cells is 1 m/s (required to achieve ~100 eps). The high quality of the images indicates that the intelligent IACS is capable of identifying the intracellular chemical distribution and morphological features of various cell types ranging from 3  $\mu\text{m}$  to 30  $\mu\text{m}$  in cell diameter. Processing of the raw images was performed using ImageJ (STAR Methods). BF, bright field; DF, dark field; Chl, chlorophyll. Scale bars, 10  $\mu\text{m}$ . (A) *Chlorella sorokiniana*, *Chlamydomonas reinhardtii*, *Haematococcus lacustris*, and *Gloeomonas anomalipyrenoides* stained with SYTO16. (B) *Euglena gracilis* stained with BODIPY<sup>505/515</sup>. (C) Single human platelets, small human platelet aggregates, and large human platelet

aggregates stained with anti-CD61 PE.(D) H1975 (adenocarcinoma) cells stained with SYTO16 and anti-EpCAM PE.(E) Erythrocytes, leukocytes, and a rare EpCAM-positive circulating cell found in fixed human blood, stained with SYTO16 and anti-EpCAM PE.

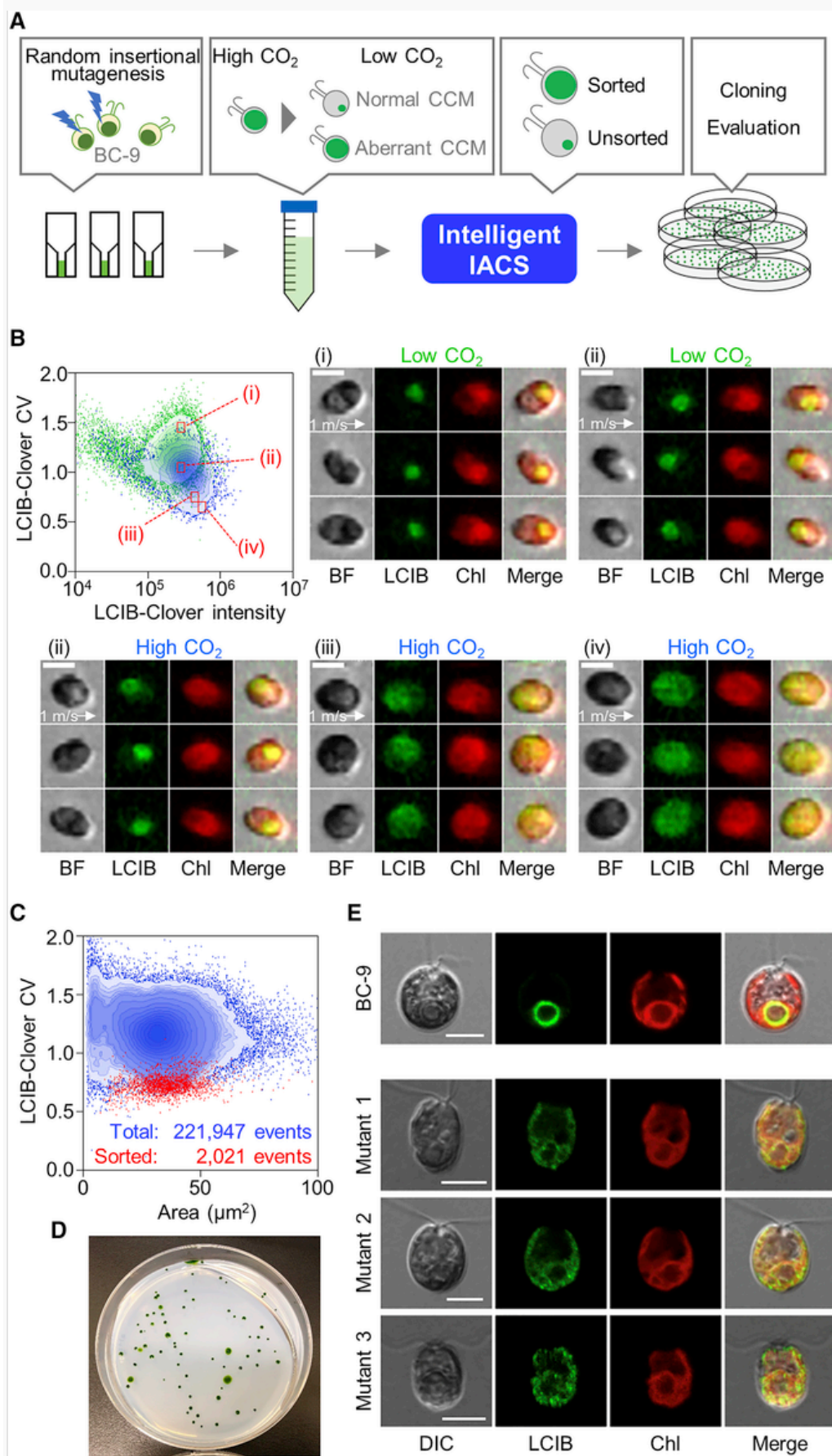
## Application to High-Content Sorting of *Chlamydomonas*

### *reinhardtii*

Ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) is a key enzyme of the Calvin-Benson cycle for photosynthetic CO<sub>2</sub> fixation. To overcome Rubisco's low level of catalytic affinity against CO<sub>2</sub> and 10,000-fold lower diffusion rate of CO<sub>2</sub> in aquatic conditions, most aquatic photosynthetic organisms possess a biophysical carbon-concentrating mechanism (CCM). In *C. reinhardtii*, the CCM is operated by active bicarbonate (HCO<sub>3</sub><sup>-</sup>) uptake regulated by a calcium-binding protein CAS (Wang et al., 2016), conversion of HCO<sub>3</sub><sup>-</sup> to CO<sub>2</sub> by carbonic anhydrases, and a non-membrane-bound sub-organelle called a pyrenoid (Freeman Rosenzweig et al., 2017; Mackinder et al., 2017), in which the CO<sub>2</sub> concentration elevates around the Rubisco. In addition to these CCM-related components, a low-CO<sub>2</sub> inducible protein B (LCIB), which is highly conserved in algae with the CCM, plays a key role in the CCM operation. In response to reduced CO<sub>2</sub> concentrations, LCIB normally changes its localization from dispersion in the chloroplast to a ring-like structure in the vicinity of the pyrenoid when the CCM is fully induced (Yamano et al., 2014). Although the algal CCM is the primary target of engineering photosynthesis in synthetic microalgae to increase their photosynthetic efficiency and yield, a full understanding of the CCM remains challenging.

We used the intelligent IACS to sort *C. reinhardtii* mutants with aberrant LCIB localization under low-CO<sub>2</sub> conditions (Figure 4A). Specifically, we used *C. reinhardtii* strain BC-9 (a transgenic line expressing LCIB fused to fluorescent protein Clover) as a sample (STAR Methods). To distinguish between LCIB localization patterns of dispersion in the chloroplast and aggregation in the vicinity of the pyrenoid, a coefficient of variation (CV) of the LCIB-Clover fluorescence signal within the cell boundary was quantified on each FDM image. As shown in Figure 4B, cells with high CV values clearly show the localization of the LCIB-Clover protein around the pyrenoid, whereas other cells with low CV values show a dispersed distribution of the protein. The intelligent IACS detected 221,947 events and sorted 2,021 low-CV cells (less than 1% of the cell population) (Figure 4C). We then placed a portion of the sorted cells (60 cells/2,021 sorted cells) onto a culture plate and obtained 44 clones (Figure 4D), out of

which 31 (70%) were recognized to possess dispersed LCIB-Clover in the chloroplast under a differential interference contrast (DIC) and confocal fluorescence microscope. Figure 4E shows the DIC and confocal fluorescence images of three such cells. A further molecular analysis of the sorted cells may lead to the identification of genes and regulatory pathways responsible for the CCM.



**Figure 4** Enrichment and Cloning of *Chlamydomonas reinhardtii* Mutants (A) Experimental procedure. Strain BC-9 was transformed by random insertional mutagenesis with a paromomycin resistance marker and cultured in acetate containing TAP medium to mimic high-CO<sub>2</sub> conditions where the CCM was inactive. Then, BC-9 cells were transferred to carbon-free TP medium to induce the CCM, followed by enrichment with the intelligent IACS. (B) Scatterplot of the cells under high-CO<sub>2</sub> (blue) and low-CO<sub>2</sub> (green) conditions in the LCIB-Clover intensity and CV together with images of the cells within the marked regions (1)–(4) in the scatterplot. (C) Scatterplot of the total population and sorted mutants. (D) Colonies of the sorted mutants in TAP agar medium containing hygromycin and paromomycin. (E) DIC and confocal fluorescence images of BC-9 and three representative mutants cultured in carbon-free TP medium. Chl, chlorophyll. Scale bars, 5  $\mu$ m.

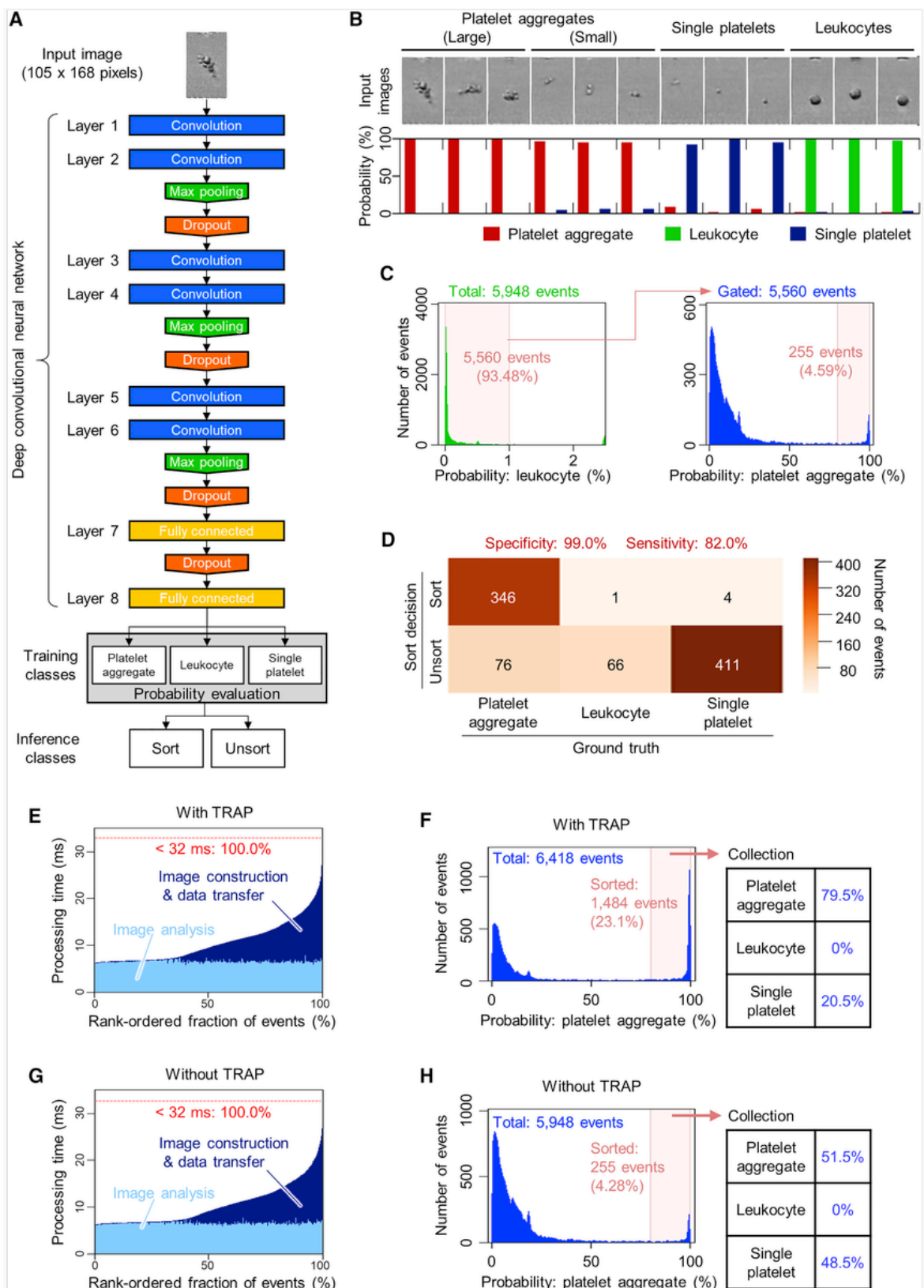
## Application to High-Content Sorting of Human Platelets

Platelet activation in circulating blood is involved in atherothrombosis including acute coronary syndromes and acute ischemic stroke (Davì and Patrono, 2007). The disorder is initiated by disruption of atherosclerotic plaques, followed by further platelet activation, resulting in local occlusion or distal embolism by resultant platelet aggregates (Libby, 2001). Evaluation of activated platelets in peripheral blood is, therefore, useful for predicting and assessing disease state. Recent findings suggest that immunothrombosis, initiated by the innate immune system, consists of neutrophils, monocytes, and platelets and contributes to arterial and/or venous thrombosis (Engelmann and Massberg, 2013). Whole-blood flow cytometry with platelet-activation marker antibodies fails to detect activated platelets accurately since platelets are sensitive to chemical and mechanical stimulation and are, hence, prone to be spontaneously activated *in vitro*, resulting in increased background noise. Fixation of blood cells is useful for minimizing the detrimental effect although it inhibits the platelet-activation marker antibodies from binding to the activated platelets. The ability to detect and isolate activated platelets (e.g., platelet aggregates) in fixed blood without staining has great potential for clinical applications, including, but not limited to, accurate diagnosis of atherothrombosis and monitoring of anti-platelet therapy.

We used the intelligent IACS to sort platelet aggregates from human blood using an eight-layer deep CNN (Figure 5A; STAR Methods). Specifically, the CNN was used to calculate the probability of identifying an object in each bright-field image as a leukocyte, single platelet, or platelet aggregate in fixed and hemolysed human blood (Figure 5B), showing the CNN's ability to differentiate even between a single platelet and a doublet. Applying progressive gating to the probabilities (Figure 5C) yielded high specificity

and sensitivity of 99.0% and 82.0%, respectively, in detecting platelet aggregates (Figure 5D). All the events were processed on the real-time intelligent image processor within 32 ms of latency (Figure 5E). As shown in Figure 5F, the intelligent IACS based on the gating strategy identified and sorted 23.1% of the cell population in a blood sample from a health donor as platelet aggregates after stimulating the sample with a thrombin receptor activator peptide (TRAP) as an agonist to activate platelets (Figure S5A; STAR Methods). Our microscopic enumeration of the sorted events verified a high sort purity of 79.5% for platelet aggregates. This high purity indicates the CNN's superior ability to enrich platelet aggregates, which is not possible with FACS or classical image analysis on the intelligent IACS, whose inability is manifested as a strong overlap between the platelet aggregates and other cells as shown in Figures S5B and S5C. As opposed to the classical image analysis, the CNN was able to identify and sort 4.28% of the events in an unstimulated blood sample (without TRAP), which simulates the *in vivo* state, out of which 51.5% were recognized as platelet aggregates by microscopic enumeration (Figures 5G and 5H). A further investigation of the sorted platelets via transcriptomic, proteomic, and metabolomic analyses may identify key factors that determine the sensitivity of platelet activation and lead to the development of a new class of anti-platelet drugs, which have been long awaited for anti-platelet therapy.





**Figure 5** Isolation and Analysis of Platelet Aggregates in Human Blood (A) CNN model architecture composed of six convolutional layers and two fully connected layers, accompanied with max-pooling and dropout connections. The final layer yields a set of probability values based on three training classes (platelet aggregate, leukocyte, and single platelet) for making sort/unsort decisions. (B) Probability distributions of representative platelet aggregate, leukocyte, and single platelet images. (C) Progressive gating strategy employed on the deep CNN to obtain platelet aggregates. (D) Correlation matrix between the ground truth and the sort decision. (E and G) Processing time of the image construction and data transfer combined and the image analysis for each event for TRAP-stimulated and unstimulated blood samples. (F and H) Histogram of gated events for TRAP-stimulated and unstimulated blood samples in the platelet-aggregate class probability with the sorting statistics. See also Figure S5.

## Discussion

We see many applications of the intelligent IACS beyond the analysis and isolation of single-cell suspensions and cell clusters we have conducted here. Although the flow channels and optical system used in the current intelligent IACS are optimized for analyzing individual cells, changing the flow channels and optical system could enable image-based sorting of larger biological objects such as cell spheroids, organoids, tissue fragments, and perhaps whole organisms (e.g., *Drosophila* embryos, zebrafish, etc.) without changing the fundamental operation of the intelligent IACS. Moreover, according to a recent report, adherent single cells can also be a subject of imaging flow cytometry experiments by growing the cells on hydrogel microcarriers (Wu et al., 2018). Importantly, these carriers possess shear-stress shelter features that prevent adherent cells from being damaged when flowing through a microchannel, suggesting compatibility for adherent-cell-based sorting. The intelligent IACS holds promise for analyzing these types of cell and tissue samples.

By virtue of the high flexibility, high scalability, and real-time automated operation of the intelligent IACS, its capabilities can further be enhanced in various directions. First, more/different imaging methods (e.g., Raman imaging, quantitative phase imaging, photoacoustic imaging), sensors, and/or CPUs/FPGAs/GPUs can simply be added as modules to the intelligent IACS for increased specificity, speed, and/or computational power without making any modifications to the intelligent IACS platform. Second, the intelligent IACS can be integrated with high-throughput single-cell -omic analysis tools in a seamless manner to streamline the identification of molecular underpinnings correlated to morphometric phenotypes. Third, the images of sorted cells can be used to further train the



deep CNN for enhanced intelligence, refine sort decisions, and optimize the accuracy of identifying extremely rare cells (e.g., circulating tumor cells, antigen-specific T cells, cancer stem cells). With these additional capabilities, potential applications of the intelligent IACS are enormous and expected to enable machine-based scientific discovery in biological, pharmaceutical, and medical sciences.

An open innovation platform is available to external researchers interested in using the intelligent IACS to test their ideas and samples. Go to the following website for more information of the open innovation platform: <http://www.goda.chem.s.u-tokyo.ac.jp/intelligentIACS>.

## Acknowledgments

This work was funded by the ImPACT program of the [Council for Science, Technology, and Innovation \(Cabinet Office, Government of Japan\)](#). We thank M. Kanematsu, M. Urakawa, and A. Komiya for assistance. N.N. is an ISAC Marylou Ingram Scholar.

## Author Contributions

K.G. conceived the concept of the intelligent IACS. H.M., T. Iino, K. Hiramatsu, C.L., Y. Suzuki, and Y. Ozeki designed and constructed the optics including the FDM microscope and speed meter and designed and implemented the hardware and software of the IC node. N.N., A.I., S.S., F.A., T.H., Y.K., S.L., M.L., Y.T., Y. Yalikun, and D.D.C. designed, fabricated, and evaluated the microfluidic chip, fluidic system, hydrodynamic and acoustic focusers, sorter, and optics-microfluidics integration unit. T.S., K. Hiraki, M.I., K.K., D.M., M.O., T.O., and H.T. designed and implemented the hardware and software of the speed meter, 10-Gbps all-IP network, IA node, TM node, and sort driver. T.E. and M. Yamada implemented the CNN. Y. Oguchi, H.S., Y. Shirasaki, N.S., and M. Yamagishi developed the centrifugation-based cell counting device. N.N., T.S., A.I., H.M., K. Hiraki, S.S., T. Iino, T.H., M.I., Y.K., K.K., D.M., M.O., H.T., Y. Hosokawa, and Y. Ozeki integrated the intelligent IACS. N.N., T.S., A.I., T. Iino, H.K., T.M., and Y. Oguchi conducted the sorting experiments and data analysis. N.N., A.I., Y. Hoshino, T. Ito, S.M., A.N., T.Y., A.Y., Y. Yatomi, M. Yazawa, Y. Hosokawa, and S.U. planned experiments. T.S., Y.F., and K.K. designed and implemented the data analysis software. H.F., M.H., T. Ito, C.T., and T.Y. prepared the microalgal cells and analyzed the results. S.M., K. Shiba, K. Suga, A.Y., and Y. Yatomi prepared the blood samples and analyzed the results. N.N., T.S., Y. Hoshino, T. Ito, A.N., Y. Hosokawa, S.U., Y. Ozeki, and K.G. supervised the work. N.N., S.M., T.Y., A.Y., Y. Yatomi, D.D.C., Y. Ozeki, and K.G. wrote the paper. A.I., H.M., and S.S. provided assistance with the figures and supplemental materials. All authors participated in editing it.

## Declaration of Interests

H.M. and K.G. are inventors of a patent application covering the FDM microscope. S.S., F.A., and T.H. are inventors of a patent application covering the dual-membrane push-pull cell sorter. N.N., T.S., and K.G. are inventors of a patent application covering the data analysis and display method. N.N. is the president of CYBO, Inc. N.N., T.S., and K.G. are shareholders of CYBO, Inc.

## Supplemental Information

Supplemental Information contains five figures and one video and can be found with this article online at <https://doi.org/10.1016/j.cell.2018.08.028>.

## References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2016). TensorFlow: large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467v2.
- Altschuler S.J. and Wu L.F., Cellular heterogeneity: do differences make a difference?, *Cell* **141**, 2010, 559–563.
- Basiji D.A., Ortyrn W.E., Liang L., Venkatachalam V. and Morrissey P., Cellular image analysis and imaging by flow cytometry, *Clin. Lab. Med.* **27**, 2007, 653, 670, viii.
- Boutros M., Heigwer F. and Laufer C., Microscopy-based high-content screening, *Cell* **163**, 2015, 1314–1325.
- Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools. Art. ID: 2236121. Available at <https://opencv.org>.
- Brasko C., Smith K., Molnar C., Farago N., Hegedus L., Balind A., Balassa T., Szkalisity A., Sukosd F., Kocsis K., et al., Intelligent image-based in situ single-cell isolation, *Nat. Commun.* **9**, 2018, 226.
- Burg T.P., Godin M., Knudsen S.M., Shen W., Carlson G., Foster J.S., Babcock K. and Manalis S.R., Weighing of biomolecules, single cells and single nanoparticles in fluid, *Nature* **446**, 2007, 1066–1069.
- Caicedo J.C., Cooper S., Heigwer F., Warchal S., Qiu P., Molnar C., Vasilevich A.S., Barry J.D., Bansal H.S., Kraus O., et al., Data-analysis strategies for image-based cell profiling, *Nat. Methods* **14**, 2017, 849–863.
- Chattopadhyay P.K., Gierahn T.M., Roederer M. and Love J.C., Single-cell technologies for monitoring immune systems, *Nat. Immunol.* **15**, 2014, 128–135.
- Chen C.L., Mahjoubfar A., Tai L.-C., Blaby I.K., Huang A., Niazi K.R. and Jalali B., Deep learning in label-free cell classification, *Sci. Rep.* **6**, 2016, 21471.
- Chen Y., Wu T.-H., Kung Y.-C., Teitell M.A. and Chiou P.-Y., 3D pulsed laser-triggered high-speed microfluidic fluorescence-activated cell sorter, *Analyst (Lond.)* **138**, 2013, 7308–7315.

- Chollet, F. (2015). Keras: Deep learning for humans. Software available from GitHub repository. Available at <https://github.com/keras-team/keras>.
- Davì G. and Patrono C., Platelet activation and atherothrombosis, *N. Engl. J. Med.* **357**, 2007, 2482–2494.
- Engelmann B. and Massberg S., Thrombosis as an intravascular effector of innate immunity, *Nat. Rev. Immunol.* **13**, 2013, 34–45.
- Eulenberg P., Köhler N., Blasi T., Filby A., Carpenter A.E., Rees P., Theis F.J. and Wolf F.A., Reconstructing cell cycle and disease progression using deep learning, *Nat. Commun.* **8**, 2017, 463.
- Fabbri F., Carloni S., Zoli W., Ulivi P., Gallerani G., Fici P., Chiadini E., Passardi A., Frassinetti G.L., Ragazzini A. and Amadori D., Detection and recovery of circulating colon cancer cells using a dielectrophoresis-based device: KRAS mutation status in pure CTCs, *Cancer Lett.* **335**, 2013, 225–231.
- Freeman Rosenzweig E.S., Xu B., Kuhn Cuellar L., Martinez-Sanchez A., Schaffer M., Strauss M., Cartwright H.N., Ronceray P., Plitzko J.M., Förster F., et al., The eukaryotic CO<sub>2</sub>-concentrating organelle is liquid-like and exhibits dynamic reorganization, *Cell* **171**, 2017, 148–162.e19.
- George T.C., Basiji D.A., Hall B.E., Lynch D.H., Ortyn W.E., Perry D.J., Seo M.J., Zimmerman C.A. and Morrissey P.J., Distinguishing modes of cell death using the ImageStream multispectral imaging flow cytometer, *Cytometry A* **59**, 2004, 237–245.
- George T.C., Fanning S.L., Fitzgerald-Bocarsly P., Medeiros R.B., Highfill S., Shimizu Y., Hall B.E., Frost K., Basiji D., Ortyn W.E., et al., Quantitative measurement of nuclear translocation events using similarity analysis of multispectral cellular images obtained in flow, *J. Immunol. Methods* **311**, 2006, 117–129.
- Girault M., Kim H., Arakawa H., Matsuura K., Odaka M., Hattori A., Terazono H. and Yasuda K., An on-chip imaging droplet-sorting system: a real-time shape recognition method to screen target cells in droplets with single cell resolution, *Sci. Rep.* **7**, 2017, 40072.
- Godin M., Delgado F.F., Son S., Grover W.H., Bryan A.K., Tzur A., Jorgensen P., Payer K., Grossman A.D., Kirschner M.W. and Manalis S.R., Using buoyant mass to measure the growth of single cells, *Nat. Methods* **7**, 2010, 387–390.
- Grys B.T., Lo D.S., Sahin N., Kraus O.Z., Morris Q., Boone C. and Andrews B.J., Machine learning and computer vision approaches for phenotypic profiling, *J. Cell Biol.* **216**, 2017, 65–71.
- Hao N., Budnik B.A., Gunawardena J. and O’Shea E.K., Tunable signal processing through modular control of transcription factor translocation, *Science* **339**, 2013, 460–464.
- Henery S., George T., Hall B., Basiji D., Ortyn W. and Morrissey P., Quantitative image based apoptotic index measurement using multispectral imaging flow cytometry: a comparison with standard photometric methods, *Apoptosis* **13**, 2008, 1054–1063.

- Hennig H., Rees P., Blasi T., Kamensky L., Hung J., Dao D., Carpenter A.E. and Filby A., An open-source solution for advanced imaging flow cytometry data analysis using machine learning, *Methods* **112**, 2017, 201–210.
- Hiraki K., Inaba M., Tezuka H., Tomari H., Koizumi K. and Kondo S., All-IP-Ethernet architecture for real-time sensor-fusion processing, *Proceedings of SPIE BiOS* **9720**, 2016, 97200D.
- Hung P.J., Lee P.J., Sabounchi P., Aghdam N., Lin R. and Lee L.P., A novel high aspect ratio microfluidic design to provide a stable and uniform microenvironment for cell growth in a high throughput mammalian cell culture array, *Lab Chip* **5**, 2005, 44–48.
- Jiang Y., Lei C., Yasumoto A., Kobayashi H., Aisaka Y., Ito T., Guo B., Nitta N., Kutsuna N., Ozeki Y., et al., Label-free detection of aggregated platelets in blood by machine-learning-aided optofluidic time-stretch microscopy, *Lab Chip* **17**, 2017, 2426–2434.
- Kasai F., Kawachi M., Erata M. and Watanabe M.M., NIES-collection. List of Strains. Microalgae and Protozoa, Seventh Edition, 2004, National Institute for Environmental Studies; Tsukuba, Japan, 257.
- Kobayashi H., Lei C., Wu Y., Mao A., Jiang Y., Guo B., Ozeki Y. and Goda K., Label-free detection of cellular drug responses by high-throughput bright-field imaging and machine learning, *Sci. Rep.* **7**, 2017, 12454.
- Lauersen K.J., Kruse O. and Mussgnug J.H., Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a versatile, modular vector toolkit, *Appl. Microbiol. Biotechnol.* **99**, 2015, 3491–3503.
- LeCun Y., Bengio Y. and Hinton G., Deep learning, *Nature* **521**, 2015, 436–444.
- LeCun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W. and Jackel L.D., Backpropagation applied to handwritten zip code recognition, *Neural Comput.* **1**, 1989, 541–551.
- Lee D., Mehta N., Shearer A. and Kastner R., A hardware accelerated system for high throughput cellular image analysis, *J. Parallel Distrib. Comput.* **113**, 2018, 167–178.
- Libby P., Current concepts of the pathogenesis of the acute coronary syndromes, *Circulation* **104**, 2001, 365–372.
- Lodha M., Schulz-Raffelt M. and Schroda M., A new assay for promoter analysis in *Chlamydomonas* reveals roles for heat shock elements and the TATA box in *HSP70A* promoter-mediated activation of transgene expression, *Eukaryot. Cell* **7**, 2008, 172–176.
- Mackinder L.C.M., Chen C., Leib R.D., Patena W., Blum S.R., Rodman M., Ramundo S., Adams C.M. and Jonikas M.C., A spatial interactome reveals the protein organization of the algal CO<sub>2</sub>-concentrating mechanism, *Cell* **171**, 2017, 133–147.e14.
- Meng P., Jacobsen M., Kimura M., Dergachev V., Anantharaman T., Requa M. and Kastner R., Hardware accelerated alignment algorithm for optical labeled genomes, *ACM Trans. Reconfig. Technol. Syst.* **9**, 2016, 10.1145/2840811.

- Mikami H., Harmon J., Kobayashi H., Hamad S., Wang Y., Iwata O., Suzuki K., Ito T., Aisaka Y., Kutsuna N., et al., Ultrafast confocal fluorescence microscopy beyond the fluorescence lifetime limit, *Optica* **5**, 2018, 117–126.
- Moor A.E., Golan M., Massasa E.E., Lemze D., Weizman T., Shenhav R., Baydatch S., Mizrahi O., Winkler R., Golani O., et al., Global mRNA polarization regulates translation efficiency in the intestinal epithelium, *Science* **357**, 2017, 1299–1303.
- Nakazawa Y., Hiraki M., Kamiya R. and Hirono M., SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole, *Curr. Biol.* **17**, 2007, 2169–2174.
- Nielsen J. and Keasling J.D., Engineering cellular metabolism, *Cell* **164**, 2016, 1185–1197.
- Ota S., Horisaki R., Kawamura Y., Ugawa M., Sato I., Hashimoto K., Kamesawa R., Setoyama K., Yamaguchi S., Fujiu K., et al., Ghost cytometry, *Science* **360**, 2018, 1246–1251.
- Pernas L., Bean C., Boothroyd J.C. and Scorrano L., Mitochondria restrict growth of the intracellular parasite *Toxoplasma gondii* by limiting its uptake of fatty acids, *Cell Metab.* **27**, 2018, 886–897.e4.
- Piatkevich K.D., Jung E.E., Straub C., Linghu C., Park D., Suk H.-J., Hochbaum D.R., Goodwin D., Pnevmatikakis E., Pak N., et al., A robotic multidimensional directed evolution approach applied to fluorescent voltage reporters, *Nat. Chem. Biol.* **14**, 2018, 352–360.
- R Core Team, R: A language and environment for statistical computing, 2016, R Foundation for Statistical Computing; Vienna, Austria, Available at <https://www.R-project.org>.
- Sakuma S., Kasai Y., Hayakawa T. and Arai F., On-chip cell sorting by high-speed local-flow control using dual membrane pumps, *Lab Chip* **17**, 2017, 2760–2767.
- Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., et al., Fiji: an open-source platform for biological-image analysis, *Nat. Methods* **9**, 2012, 676–682.
- Stubbington M.J.T., Rozenblatt-Rosen O., Regev A. and Teichmann S.A., Single-cell transcriptomics to explore the immune system in health and disease, *Science* **358**, 2017, 58–63.
- Sun T., Kovac J. and Voldman J., Image-based single-cell sorting via dual-photopolymerized microwell arrays, *Anal. Chem.* **86**, 2014, 977–981.
- Tung J.W., Heydari K., Tirouvanziam R., Sahaf B., Parks D.R., Herzenberg L.A. and Herzenberg L.A., Modern flow cytometry: a practical approach, *Clin. Lab. Med.* **27**, 2007, 453, 468, v.
- von Erlach T.C., Bertazzo S., Wozniak M.A., Horejs C.-M., Maynard S.A., Attwood S., Robinson B.K., Autefage H., Kallepitis C., Del Río Hernández A., et al., Cell-geometry-dependent changes in plasma membrane order direct stem cell signalling and fate, *Nat. Mater.* **17**, 2018, 237–242.

- Wakisaka Y., Suzuki Y., Iwata O., Nakashima A., Ito T., Hirose M., Domon R., Sugawara M., Tsumura N., Watarai H., et al., Probing the metabolic heterogeneity of live *Euglena gracilis* with stimulated Raman scattering microscopy, *Nat. Microbiol.* **1**, 2016, 16124.
- Wang L., Yamano T., Takane S., Niikawa Y., Toyokawa C., Ozawa S.-I., Tokutsu R., Takahashi Y., Minagawa J., Kanesaki Y., et al., Chloroplast-mediated regulation of CO<sub>2</sub>-concentrating mechanism by Ca<sup>2+</sup>-binding protein CAS in the green alga *Chlamydomonas reinhardtii*, *Proc. Natl. Acad. Sci. USA* **113**, 2016, 12586–12591.
- Wu C.-Y., Stoecklein D., Kommajosula A., Lin J., Owsley K., Ganapathysubramanian B. and Di Carlo D., Shaped 3D microcarriers for adherent cell culture and analysis, *Microsys. Nanoeng.* **4**, 2018, 21.
- Yamano T., Asada A., Sato E. and Fukuzawa H., Isolation and characterization of mutants defective in the localization of LCIB, an essential factor for the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*, *Photosynth. Res.* **121**, 2014, 193–200.
- Yamano T., Iguchi H. and Fukuzawa H., Rapid transformation of *Chlamydomonas reinhardtii* without cell-wall removal, *J. Biosci. Bioeng.* **115**, 2013, 691–694.
- Zenker J., White M.D., Templin R.M., Parton R.G., Thorn-Seshold O., Bissiere S. and Plachta N., A microtubule-organizing center directing intracellular transport in the early mouse embryo, *Science* **357**, 2017, 925–928.

**Keywords:** image-activated cell sorting; high-throughput microscopy; high-throughput screening; deep learning; convolutional neural network; cellular heterogeneity; machine intelligence; cellular morphology

## STAR★Methods

### Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Mouse monoclonal anti-CD61, PE conjugated	Beckman Coulter	Cat#IM3605
Mouse monoclonal anti-EpCAM (clone VU-1D9)	GeneTex	Cat#GTX28667
Rat monoclonal anti-mouse IgG1 (clone A85-1), PE conjugated	BD Biosciences	Cat#550083
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
RPMI medium 1640	Thermo Fisher Scientific (GIBCO)	Cat#11875093
Fetal bovine serum	Thermo Fisher Scientific (GIBCO)	Cat#10270-106
SYTO16	Thermo Fisher Scientific (Invitrogen)	Cat#S7578
Glutaraldehyde (25% solution)	Nacalai Tesque, Kyoto, Japan	Cat#17003-92
TRAP-6 amide trifluoroacetate salt	BACHEM	Cat#H-2936.0005

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hygromycin	Wako	Cat#085-06153
Paromomycin	Wako	Cat#161-23603
Bovine serum albumin solution	Sigma	Cat#A9576
<b>Critical Commercial Assays</b>		
EasyComp fluorescent particle kit (Blank, FITC, PE & PE-Cy5), 4 vials	Spherotech	Cat#ECFP-K1
Rainbow fluorescent particles, 6.0-6.4 $\mu$ m	Spherotech	Cat#RFP-60-5
Rainbow calibration particles, 6 peaks	Spherotech	Cat#RCP-30-5 (6 peaks)
Polystyrene particle size standard kit, flow cytometry grade	Spherotech	Cat#PPS-6K
Lysing buffer (10X concentrated), BD pharm lyse	BD Biosciences	Cat#555899
OptiLyse C	Beckman Coulter	Cat#A11895
PrimeSTAR GXL DNA polymerase	Takara Bio	Cat#R050A
In-Fusion HD cloning kit	Takara Bio	Cat#639648
MAX efficiency transformation reagent for algae	Thermo Fisher Scientific (Invitrogen)	Cat#A24229
<b>Experimental Models: Cell Lines</b>		
Human adenocarcinoma: H1975 cells	ATCC	Cat#CRL-5908
<b>Experimental Models: Organisms/Strains</b>		
<i>Euglena gracilis</i>	Microbial Culture Collection at NIES	Cat#NIES-48



REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Chlorella sorokiniana</i>	Tsuruoka, Keio, Algae Collection	Cat#TKAC1027
<i>Chlamydomonas reinhardtii</i>	Tsuruoka, Keio, Algae Collection	Cat#TKAC1017
<i>Chlamydomonas reinhardtii</i> C-9	Microbial Culture Collection at NIES	Cat#NIES-2235
<i>Chlamydomonas reinhardtii</i> BC-9	This paper	N/A
<i>Haematococcus lacustris</i>	Microbial Culture Collection at NIES	Cat#NIES-4141
<i>Gloeomonas anomalipyrenoides</i>	Microbial Culture Collection at NIES	Cat#NIES-3640
<b>Oligonucleotides</b>		
Primer: LCIB-N-fusion-F (forward): TTTGCAGGATGCATATGTTTCGCTCTGTCTTCGCGC	This paper	N/A
Primer: LCIB-N-fusion-R (reverse): CGATGACGTCAGATCTGTTCTTGGGGGCCTCGAA	This paper	N/A
Primer: aph-F1 (forward): GCTTATCGATACCGTCGACCT	This paper	N/A
Primer: aph-R3 (reverse): AACACCATCAGGTCCTCAG	This paper	N/A
<b>Recombinant DNA</b>		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pOpt_Clover_Hyg vector	Lauersen et al., 2015; GenBank	Cat#KM061067.2
LCIB-Cloverexpression plasmid	This paper	N/A
pGenD-aphVIII	Nakazawa et al., 2007, personally provided	N/A
<b>Software and Algorithms</b>		
ImageJ	NIH	<a href="https://imagej.nih.gov/ij">https://imagej.nih.gov/ij</a>
TensorFlow	Abadi et al., 2016	<a href="https://arxiv.org/abs/1603.04467">https://arxiv.org/abs/1603.04467</a>
Keras	Chollet, 2015	<a href="https://github.com/keras-team/keras">https://github.com/keras-team/keras</a>
R	R Core Team, 2016	<a href="https://www.R-project.org/">https://www.R-project.org/</a>
OpenCV	Bradski, 2000	<a href="https://opencv.org/">https://opencv.org/</a>
LabVIEW	National Instruments	<a href="http://www.ni.com/labview/">http://www.ni.com/labview/</a>

## Contact for Reagent and Resource Sharing

Further information and requests for reagents and resources should be directed to and will be fulfilled by the Lead Contact, Keisuke Goda ([goda@chem.s.u-tokyo.ac.jp](mailto:goda@chem.s.u-tokyo.ac.jp)).

## Experiemntal Model and Subject Details

### Calibration particles

To evaluate the sorting performance of the intelligent IACS, a 1:1 mixture and a 1:19 mixture of 3- $\mu$ m particles (3.0-3.4  $\mu$ m PE-Cy5 particles included in EasyComp fluorescent particle kit, Spherotech Inc., IL, USA) and 6- $\mu$ m particles (Rainbow fluorescent particles, 6.0-6.4  $\mu$ m, Spherotech Inc., IL, USA) were injected from the sample inlet into the microfluidic chip. To evaluate the imaging performance of

the FDM microscope, standard 6-peak particles (Rainbow calibration particles, 6 peaks, RCP-30-5, Spherotech Inc.), a standard compensation particle kit (EasyComp fluorescent particle kit, Spherotech, Inc.), and a size standard particle kit (Polystyrene particle size standard kit, flow cytometry grade, PPS-6K, Spherotech Inc.) were employed. The concentration of the particles was tuned to be  $1 \times 10^6$  particles/mL, which corresponds to  $\sim 100$  eps in the intelligent IACS.

### Cultured cells

Fresh water and marine microalgae which express diverse cell size and morphology, *E. gracilis* NIES-48, *C. sorokiniana* TKAC1027, *C. reinhardtii* TKAC1017 (NIES-2463), *G. anomalopyrenoides* NIES-3640, and *H. lacustris* NIES-4141 were provided by Microbial Culture Collection at the National Institute for Environmental Studies (NIES) (<http://mcc.nies.go.jp/>) and Tsuruoka, Keio, Algae Collection (TKAC) of T. Nakada at Institute for Advanced Biosciences, Keio University. The cultures were grown in culture flasks (working volume: 20 ml) under 14:10 light:dark cycle with illumination at 120-140  $\mu\text{mol photons/m}^2/\text{s}$  at 25°C. The culture of each algal strain was grown in AF-6 medium (Kasai et al., 2004). A human cell line H1975 (adenocarcinoma, female) was obtained from ATCC and maintained in RPMI medium 1640 (Thermo Fisher Scientific) with 10% fetal bovine serum (Thermo Fisher Scientific) in a humidified atmosphere of 5% CO<sub>2</sub> at 37°C. The H1975 cells were collected using 0.01% EDTA and 0.1% trypsin, followed by washing with PBS and labeling using SYTO16 and anti-EpCAM antibody labeled with PE-conjugated second antibody as described above.

### Blood samples for detection of EpCAM-positive circulating cells

We collected blood samples from healthy volunteers and patients seen at the Cancer Institute Hospital of the Japanese Foundation for Cancer Research (JFCR). A whole-blood sample (5 ml) was collected from a patient with pancreatic cancer. Blood samples were filtered using a size-selective microfilter (10- $\mu\text{m}$  pore V-type, Optnics Precision Co. Ltd., Ashikaga, Japan). The cells trapped on the filter were washed with 10 mL of PBS containing 2 mM EDTA (PBS-E), treated with 10x-diluted lysing buffer BD pharm lyse (BD Biosciences, San Jose, CA), washed with 10 mL of PBS-E, and mechanically recovered from the filter with 600  $\mu\text{L}$  of PBS-E. The nuclei of the collected cells were stained by adding 0.6  $\mu\text{L}$  of SYTO16 (Thermo Fisher Scientific, Waltham, MA) for 10 min at ambient temperature and fixed with 0.25% of glutaraldehyde (Nacalai tesque, Kyoto, Japan) for 10 min at ambient temperature, followed by washing with 5 mL of PBS-E. The fixed cells were collected by centrifuging 600 g for 5 min (5500, Kubota, Tokyo, Japan) and resuspended with 1 mL of PBS-E containing 10  $\mu\text{L}$  of anti-EpCAM antibody (VU-1D9, GeneTex, Irvine, CA). After 30 min of incubation at ambient temperature, the cells were washed with 1 mL of PBS-E twice and reacted with 2.5  $\mu\text{L}$  of PE-conjugated rat anti-mouse IgG1 (A85-

1, BD Bioscience) for 20 min at ambient temperature, followed by washing with 1 mL of PBS-E twice. The resultant stained cells were resuspended in 1 mL of PBS containing 0.05% Tween-20 and stored at 4°C with light blocking until used. The patient was 71-year-old male with histologically confirmed pancreatic adenocarcinoma. This study was approved by the Institutional Ethics Committee of the Japanese Foundation for Cancer Research (No. 2017-1096), which was conducted according to the Declaration of Helsinki. All healthy volunteers and patients provided written informed consent.

### **Blood samples for detection of platelet aggregates**

We prepared blood samples from four healthy volunteers (23-year-old male, 38-year-old male, 39-year-old male, and 50-year-old male) for training and validating the convolutional neural network (CNN) and from a 69-year-old healthy female for the sorting experiments. The blood samples were collected via a 21-gauge butterfly needle into 4.5-mL vacuum plasma separator tubes with 0.5-mL of 3.2% sodium citrate after filling the tube with blood. Next, 100  $\mu$ L of citrated blood was gently added to a tube in which 20  $\mu$ L of CD61-PE was added in advance in absence of light, followed by adding 1  $\mu$ L of 1  $\mu$ M TRAP-6 amide trifluoroacetate salt (TRAP; BACHEM) or PBS (without TRAP) and mixing thoroughly. Finally, 500  $\mu$ L of OptiLyse C (Beckman Coulter), a hemolyzing and fixing solution containing 1.5% formaldehyde, was added and mixed gently and slowly by inversion. After standing for 10 min, PBS (-) was added and mixed likewise. Each blood sample was diluted with PBS to maintain a cell concentration of  $10^7$  cells/mL and tested with the intelligent IACS within 5 hours after the sample preparation to ensure the platelet aggregates. This procedure was performed at room temperature (18 – 25°C). This study was approved by the Institutional Ethics Committee of Faculty of Medicine, the University of Tokyo (No. 11049-5) and conducted according to the Declaration of Helsinki. All volunteers provided written informed consent.

## **Method Details**

### **Principles of the FDM microscope**

We employed FDM imaging for continuous high-speed blur-free bright-field and two-color fluorescence image acquisition of flowing cells at 1 m/s. The configuration of the microscope is a modified version of the FDM microscope described in a previous report (Mikami et al., 2018). The principles of the microscope are described as follows. A pair of acousto-optic deflectors (AODs) driven by multitone electrical signals are used to convert a 488-nm continuous-wave laser beam into multiple intensity-modulated excitation beams with distinct modulation frequencies and propagation directions. By focusing these beams with an objective lens (Leica, HC PL APO CS2, 20x, NA = 0.75), a linear array of focal spots is formed in the microfluidic channel. The microfluidic chip is aligned to the microscope so that cells flow

at a fixed speed along the direction orthogonal to the linear beam array. The beam spots scan the cells in the flow direction to acquire one-dimensional (1D) spatial profiles of the transmitted excitation beams and the fluorescence signals from each cell. The transmitted excitation beams and two-color fluorescence signals [508 – 580 nm (FL1) and > 580 nm (FL2)] are separately detected by single-pixel photodetectors. The frequency range of the excitation beam array is 4 – 200 MHz. The signal waveform of each photodetector is digitized by a digitizer at a sampling rate of 1 GS/s (Spectrum M4i.2212-x8) embedded in the IC node. The 1D spatial profiles contained in the digitized signal waveforms are separated by taking their Fourier transform in the frequency domain since each beam spot has the distinct modulation frequency as mentioned above. Finally, three-color images of each cell are obtained by stacking the 1D spatial profiles in the flow direction. Consequently, the FDM microscope can obtain blur-free bright-field and fluorescence images of flowing cells at 1 m/s with a pixel size of 0.25  $\mu\text{m}$  (in the flow direction)  $\times$  0.84  $\mu\text{m}$  (in the spot array direction) and the 1D field of view of 42  $\mu\text{m}$  in the spot array direction. The total signal processing time for individual events differs, depending on the cell-cell interval in the flow. The software in the IC node is coded with LabVIEW (National Instruments).

### **Evaluation of the FDM microscope**

The imaging performance of the FDM microscope was evaluated with fluorescent particles. First, the detection sensitivity, dynamic range, and linearity of the FDM microscope were characterized by using standard 6-peak particles as shown in Figures S3E and S3F, indicating its ability to distinguish fluorescent particles with different fluorescence intensities with a sensitivity of 1,106 and 996 molecules of equivalent soluble fluorochrome (MESF) in FL1 and FL2, respectively, a dynamic range of  $\sim 100$ , and a linearity of  $> 0.99$ . Second, the spectral resolution to distinguish different fluorescence wavelengths was evaluated with FITC, PE, PE-Cy5, and Blank fluorescent calibration particles. As shown in Figure S3G, a population of particles of each species is plotted at different locations in the FL1 versus FL2 plot as expected. Finally, the spatial resolution was evaluated with size standard particles, showing a good correlation between the particle diameter squared and the calculated cross-sectional area of the particle from 2.0  $\mu\text{m}$  to 10.1  $\mu\text{m}$  (Figure S3H). These results demonstrate the FDM microscope is capable of generating images adaptable for quantitative image analysis on the intelligent IACS.

### **Architecture of the real-time intelligent image processor**

As shown in Figures 1 and S2, the real-time intelligent image processor consists of four key components: (i) the TM node which controls the timing of the image acquisition and sorting, (ii) the IC node which decodes the incoming FDM signals, (iii) the two IA nodes which analyze the images to make a sort/unsort decision for every event, and (iv) the all-IP network to combine all the components via the

network switch to operate in real time. Details of each component are as follows. The TM node is on an FPGA board (Xilinx KC705) equipped with a 1-Gbps Ethernet port. It manages the timing of the entire process including the cell detection, FDM image acquisition, and sorting based on the speed meter signals (the details of which are described below). The IC node is equipped with two multi-core CPUs, 10-Gbps Ethernet ports, and a 4-ch high-speed digitizer (Spectrum M4i.2212-x8) operating at a sampling rate of 1 GS/s, which are connected by a PCI-express bus. The outputs of the photodiodes in the FDM microscope are connected to Ch. 0 – Ch. 2 of the digitizer, while the serial digital signal from the TM node is connected to Ch. 3. In the IC node, cell images are constructed, encapsulated in UDP packets, and sent to one of the IA nodes. The digitizer is triggered by the electrical signal from the TM node while the image processing is implemented as software coded with LabVIEW (National Instruments). IA node 1 is equipped with a multi-core CPU (Intel Core i7-990X) while IA node 2 is equipped with a multi-core CPU (Core i7-8700K) and a GPU (NVIDIA GTX-1050Ti). Both of the IA nodes have two SSD storages and 10-Gbps Ethernet ports. One of the IA nodes makes a sort/unsort decision based on the results of the image analysis. In parallel, the IA node stores all the information of the processed cell images on the SSDs. The image analysis at the IA nodes is implemented as software, allowing us to employ complex algorithms including the deep CNN with highly parallel processing on the GPU. The all-IP network is a general-purpose interconnection of a 10-Gbps Ethernet and a 1-Gbps Ethernet. We used an off-the-shelf Ethernet switch (NETGEAR XS716T-100AJS) which can accommodate Ethernet connections at various bit rates. The all-IP network makes it easy to change the system configuration with high scalability simply by arranging the computing resources such as the number of CPUs on the IC/IA nodes, depending on the required processing load such as the event rate of the intelligent IACS. It is important to note that the infrastructure allows us to expand the system with high flexibility by connecting additional/other imaging methods and/or sensors to the all-IP network.

### **Signal processing in the real-time intelligent image processor**

In the real-time intelligent image processor (Figures 1 and S2), the TM/IC/IA nodes conduct various types of signal processing including time management, FDM image construction, image analysis, and sort/unsort decision making. Details of the signal processing flow are described as follows. When a cell passes through the optical interrogation points OI1 and OI3, the corresponding photodiode generates an electrical pulse whose waveform is digitized and analyzed by the TM node. Specifically, when the OI1 signal exceeds a specified threshold value, the TM node recognizes an event, assigns a cell ID number to the event, and gives a time stamp as a passage time. After a predetermined duration (typically 160  $\mu$ s), the cell ID signal and the trigger signal are sent to the IC node as two-channel electrical serial signals. After

the OI3 signal is detected, the TM node calculates the speed and estimates the sort time. All the calculated information about each event is temporarily stored in a memory on the FPGA until the corresponding sort time. In parallel, the cell ID and the speed value are sent to one of the IA nodes via the UDP communication. When the digitizer in the IC node receives the trigger signal, the IC node begins the acquisition of the FDM signals. According to the algorithm described above, three-channel FDM images are produced with a typical field of view of  $42\ \mu\text{m} \times 26.25\ \mu\text{m}$  and an image size of  $50\ \text{pixels} \times 105\ \text{pixels}$  with 16-bit depth. The number of the 1D spatial profiles in the flow direction varies, depending on the type of cells. The cell ID is received by the Ch. 3 of the digitizer. The image data and the cell ID are encapsulated in UDP packets and sent to one of the IA nodes. The IA node begins image analysis when it receives all the UDP packets of the images and cell ID from the IC node and those of the speed information from the TM node. If some packets are not delivered within a specified time, the event is aborted. The image analysis algorithm is coded with OpenCV on C++. First, the pixel size is normalized to a square, and consequently, the image size becomes  $168\ \text{pixels} \times 105\ \text{pixels}$ . A sort decision is made based on the classical image analysis or the deep CNN. The classical image analysis is used for the decision in sorting particles, *C. reinhardtii* cells, and platelets, whereas the deep CNN is used for the decision in sorting particles and platelets, the details of which are described below. It is also possible to implement other image analysis algorithms as long as the total signal processing time fits within the sort latency. Finally, the sort/unsort decision signal is sent from the IA node to the TM node with the cell ID via the UDP communication. When the TM node receives a sort decision, a sort trigger signal is transmitted to the sort driver at the sort time of the corresponding ID. In the sort driver, a predetermined signal waveform is generated and sent through a DAC to actuate the dual-membrane push-pull cell sorter. While similar to conventional imaging flow cytometers (George et al., 2006; Caicedo et al., 2017; Basiji et al., 2007; Lee et al., 2018; George et al., 2004; Henery et al., 2008), most of the raw parameters extracted through the image analysis (e.g., circularity, perimeter, nucleus-to-cytoplasm ratio) are not easy to interpret in a biological context. To handle such highly multidimensional data, various machine learning algorithms (Caicedo et al., 2017; Grys et al., 2017; Hennig et al., 2017) have been adopted, including support vector machines for platelet detection (Jiang et al., 2017) and drug screening (Kobayashi et al., 2017) and deep learning for cell cycle analysis (Eulenberg et al., 2017) and cancer cell classification (Chen et al., 2016). The flexible software-based real-time machine learning approach is an ideal architecture to apply these algorithms to real-time sorting operation.

### **Design and fabrication of the microfluidic chip**



The microfluidic chip plays the following important roles in the intelligent IACS: (i) continuously introducing cells to the FDM microscope, (ii) delivering them from the FDM microscope to the sort point with a predictable latency, (iii) guiding the sorted/unsorted cells to the collection/waste outlets. To meet these requirements, we designed the microfluidic chip as shown in Figures 1 and S1A. The microfluidic chip was fabricated using micro-electromechanical systems (MEMS) techniques (Sakuma et al., 2017; Burg et al., 2007; Godin et al., 2010; Hung et al., 2005). Since the length of the microfluidic chip is as large as a silicon wafer, we fabricated the microfluidic chip with a modified fabrication protocol to the protocol described in a previous report (Sakuma et al., 2017). As shown in Figure S1A, the microfluidic chip has three layers: a base layer, a microchannel layer, and a cover layer. To form rigid microchannels that avoid potential instability due to flow-induced deformation of microchannels in the chip, we employed 200- $\mu\text{m}$ -thick borosilicate glass substrates for the base and cover layers and a 200- $\mu\text{m}$ -thick silicon substrate for the microchannel layer. The use of such thin glass substrates is important to avoid physical interference with a pair of high-NA objective lenses. The microchannels including the inlets, outlets, and on-chip dual membrane pumps were patterned using deep reactive ion etching (DRIE). The cross-sectional dimensions of the central microchannel including the four optical interrogation points (OI1 – OI4) and the sort point are 200  $\mu\text{m} \times 200 \mu\text{m}$ . The details of the fabrication process are described in a previous report (Sakuma et al., 2017). First, to reduce the local stress concentration in the bonding process, we formed a grating pattern on the borosilicate glass surface of the base and cover layers by using the dry etching technique. Second, to obtain the masks for the wet etching process of the cover and base layers, we patterned the sputtered chromium (Cr) and gold (Au) layers and the spin-coated positive photoresist OFPR (Nihon Kayaku Co., Ltd., Japan). Third, we patterned microchannels on the base and cover layers using a wet-etching process with a depth and width of 100  $\mu\text{m}$  and 400  $\mu\text{m}$ , respectively. The etching masks were then removed. Fourth, the patterned base layer and a silicon (Si) substrate were bonded by anodic bonding. Fifth, the Si layer was etched by the DRIE process. Sixth, the etching mask for the sandblasting process of the cover layer was patterned by applying SCM250 (Nikko-Materials Co., Ltd., Japan) to the back side of the wet etched surface. Seventh, the inlets and outlets were formed by sandblasting the cover layer through the etching mask. Finally, the microfluidic chip was obtained by packaging the microchannel and the cover layer using the anodic bonding.

#### **Optics-microfluidics integration unit**

To precisely position the central microchannel for the FDM microscope and speed meter, we developed an optics-microfluidics integration unit. It consists mainly of three objective lenses and the microfluidic chip integrated with a chip holder as shown in Figure S1C. Two objective lenses A1 and A2



are aligned to image the cells flowing at the center of the microchannel while the other objective lens B is used for monitoring the sorting process with the high-speed CMOS camera. The three detection spots of the speed meter (OI1, OI3, OI4) also share these objective lenses with the FDM microscope. The microfluidic chip was sandwiched between two holding plates of the chip holder with tube connectors and piezoelectric actuators and then inserted into a slot in the integration unit (Figure S1D). Details of the installation and optical interrogation (OI1 – OI4) are shown in Figure S1D.

### **Principles of the on-chip dual-membrane push-pull cell sorter**

The microfluidic chip contains dual membrane pumps for the fast push-pull sorting actuation. The cell sorter is based on rapidly controlling local flow at the sort point and isolating target cells from the central stream with piezoelectrically actuated dual membrane pumps. The configuration of the cell sorter is a modified version of the previously reported cell sorter (Sakuma et al., 2017). Each external piezoelectric actuator was set on the corresponding glass membrane fabricated as a part of the microfluidic chip shown in Figure S1A. Since the glass membrane is deformed by the motion of the piezoelectric actuator, the local flow can be produced in the direction perpendicular to the cell flow. When the dual membrane pumps are actuated out of phase, high-speed local-flow crossing the main microchannel is produced at the sort point as shown in Figure S1B. When the pump is turned off, cells flow into the central branch of the three-branch microchannel junction, whereas when the pump is turned on at the right timing, cells flow into either the upper or lower branch (Figure S1B). The upper and lower channels are connected at the downstream and share one outlet. Therefore, sorted cells are collected from the outlet regardless of the direction of the local flow. This configuration is useful for high-throughput sorting because the initialization of the dual membrane pumps is not necessary. For sorting, we applied a ramp voltage signal with an amplitude of 80 V and a rise time of 200  $\mu$ s. The time-resolved sorting performance is shown in Figure 2C.

### **Hydrodynamic focusing**

The microfluidic chip is based on a two-step 3D hydrodynamic focusing technique (Sakuma et al., 2017; Chen et al., 2013) to align cells at the center of the microchannel before the FDM microscope (Figures 1 and S1A). Since the flow speed distribution in the microchannel has a parabolic profile (i.e., the speed is maximum at the center), the position of cells in the cross section of the microchannel greatly affects not only the yield of image detection, but also the recovery of sorting. To obtain a reproducible and stable flow speed for cells, an ideal approach is to tightly focus the cells at the center of the microchannel. The two-step 3D hydrodynamic focuser meets this purpose. As shown in Figure S1A, the two-step 3D hydrodynamic focuser uses a sheath flow to confine flowing cells in the horizontal direction, followed by

another sheath flow to confine them in the vertical direction. One of the major features of this focuser is that it can be embedded inside a thin microfluidic chip. Because the precise control of the flow rate ratio is essentially important for stable hydrodynamic focusing, we developed a flow control system which consists of pressure vessels, electro-pneumatic regulators, and liquid flow meters (Sensirion AG Switzerland). The regulators are controlled in real time based on a flow sensor measurement. With the flow control system, we experimentally evaluated the focusing performance using 6- $\mu\text{m}$  particles by analyzing images obtained with the high-speed CMOS camera V1211 (Vision Research Inc., NJ, USA). We measured the position and speed of the flowing particles, as well as the latency from the FDM microscope interrogation point to the sort point as shown in Figures S3A–S3D. The particles were tightly focused to the center of the microchannel under the condition of a flow rate ratio of sample:sheath1:sheath2 = 1:16:40. Moreover, the tight focusing served to minimize the effect of the parabolic flow speed distribution, ensuring a predictable sort latency.

### Acoustic focusing

In addition to hydrodynamic focusing, 3D acoustic focusing can optionally be used to keep flowing cells at the center of the microchannel all the way from the optical interrogation region of the FDM microscope to the sort point (Figures 1 and S1A) when needed, especially for large cells. In other words, acoustic focusing serves to avoid unwanted position shifts of flowing cells due to inertial lift and gravitational forces which would otherwise introduce cell-morphology-dependent fluctuations in the latency which is more evident for large cells. A pair of piezoelectric transducers (Fuji Ceramics Corporation, 3.66Z20\*20S-SYX) and a 0.57-mm thick 20 mm x 20 mm element were glued on both glass substrates of the microfluidic chip with an epoxy resin (7004, 3M Japan Limited, Tokyo, JAPAN). Acoustic focusing was accomplished by exciting both vertical and horizontal resonance modes of the 200  $\mu\text{m}$   $\times$  200  $\mu\text{m}$  cross section of the microchannel by actuating the piezoelectric transducers with a sinusoidal driving signal at 3.66 MHz and 142  $V_{\text{pp}}$  provided by a function generator (NF Corporation, WF1974) via a high-voltage amplifier (NF Corporation, HSA4101). We evaluated the performance of the acoustic focusing using *E. gracilis* cells at the sort point using the high-speed CMOS camera under the condition of hydrodynamic focusing with flow ratios of sample:sheath1:sheath2 = 1:16:40. Without the acoustic focusing, the *E. gracilis* cells were widely distributed in the microchannel (Figure S3C). With the acoustic focusing, the cells were tightly focused at the center of the microchannel, showing a well-focused speed profile (Figure S3D).

### Precise sort timing control

An important element of the intelligent IACS is the precise control of the sort timing to accurately isolate target cells from a single continuous stream of cells with a cell speed meter (Figure 1). While in FACS it typically takes only a few hundred microseconds for a cell to flow from the optical interrogation point to the sort point, the sort latency of the intelligent IACS needs to be  $> 100$  times longer than that of FACS due to its requirement for analyzing the much larger amount of data (images) on the FPGA-CPU-GPU system (Lee et al., 2018; Meng et al., 2016) in real time. To achieve the optimum sort timing, the intelligent IACS uses a speed meter to measure the flow speed of each cell at the FDM microscope (Figure 1) and hence to predict the latency at the sorter with a high precision of  $200\ \mu\text{s}$  (Figures S4C and S4D). As illustrated in Figure S4A, to cope with residual speed fluctuations of cells, the intelligent IACS has two optical interrogation spots (OI1 and OI3) adjacent to the FDM microscope (OI2) and one beam spot slightly upstream of the sorter (OI4). Every optical interrogation spot has a size of approximately  $6\ \mu\text{m} \times 80\ \mu\text{m}$  in the flow direction and its perpendicular direction, respectively. When cells pass through each optical interrogation spot, forward scattering (FSC) occurs and is detected by a photodetector. Figure S4B shows an example of the FSC signals detected at OI1, OI3, and OI4. The FSC signals at these spots are digitized and sent to the TM node to calculate the speed as described above. When the signal intensity at the OI1 spot exceeds a predetermined threshold, the TM node generates a trigger signal for the FDM microscope to initiate bright-field and fluorescence imaging as well as for the TM node internally to initiate the speed meter function to predict the sort latency. The predicted latency is recorded in the TM node along with a unique ID number for the event. The ID number is transferred to the IC node with the trigger signal and then transferred to the IA node along with the bright-field and fluorescence images of each event. After the images are analyzed and a sort/unsort decision is made, the ID number is transferred back to the TM node. The TM node uses the received ID number to find the latency prediction of each event and then sends a sort trigger signal to the sort driver at a calculated timing.

### Plasmid construction and isolation of transgenic lines expressing LCIB-Clover

The genomic sequence of *LCIB* was amplified by PCR with PrimeSTAR GXL DNA polymerase (Takara Bio) using genomic DNA extracted from *C. reinhardtii* strain C-9 (photosynthetically wild-type strain available from National Institute for Environmental Studies (NIES) as strain NIES-2235) as a template with forward primer LCIB-N-fusion-F (5'-TTTGCAGGATGCATATGTTTCGCTCTGTCTTCGCGC-3') and reverse primer LCIB-N-fusion-R (5'-CGATGACGTCAGATCTGTTCTTGCGGGCCTCGAA-3'). The PCR products were purified and cloned into pOpt\_Clover\_Hyg vector (Lauersen et al., 2015) digesting with restriction enzymes *NdeI* and *BglII* using an In-Fusion HD cloning kit (Takara Bio). This expression plasmid of LCIB-Clover was

transformed into the C-9 cells by electroporation using a NEPA-21 electroporator (NEPAGENE), as described in a previous report (Yamano et al., 2013). The transformants were incubated at 25°C for 24 h with gentle shaking and illumination at 1.5  $\mu\text{mol photons/m}^2/\text{s}$  and spread over Tris-acetate-phosphate (TAP) medium (TAP medium, <https://www.chlamycollection.org/methods/media-recipes/tap-and-tris-minimal/>) agar plates containing 30  $\mu\text{g/mL}$  hygromycin. Among 288 hygromycin resistant transformants, 21 lines designated as BC-1 through BC-21 showing fluorescence derived from LCIB-Clover in the vicinity of the pyrenoid were screened. Strain BC-9 showing the strongest fluorescence signal was selected for further experiments.

### Random insertional mutagenesis of strain BC-9 and preparation for cell sorting

A 1,534-bp DNA fragment containing the aminoglycoside 3'-phosphotransferase (*aphVIII*, paromomycin-resistance gene) driven by a *HSP70A-RBCS2* tandem promoter (Lodha et al., 2008) was amplified by PCR from plasmid pGenD-aphVIII (Nakazawa et al., 2007) using PrimeSTAR GXL DNA Polymerase with forward primer aph-F1 (5'-GCTTATCGATACCGTCGACCT-3') and reverse primer aph-R3 (5'-AACACCATCAGGTCCCTCAG-3'). The PCR product was purified and the concentration was adjusted to 100 ng/ $\mu\text{L}$ . This PCR product of *aphVIII* was randomly inserted into the genome of strain BC-9 as a DNA marker by electroporation as described in a previous report (Yamano et al., 2013) with a slight modification. In brief, BC-9 cells were cultured at 25°C until the cell density reached  $1\text{--}2 \times 10^6$  cells/mL. The cultured cells were collected and resuspended in a MAX efficiency transformation reagent for algae (Thermo Fisher Scientific) to a final density of  $1 \times 10^8$  cells/mL. Then, 3  $\mu\text{L}$  of 100 ng/ $\mu\text{L}$  *aphVIII* PCR product was added to 117  $\mu\text{L}$  of the cell suspension. The cell suspension was placed into an electroporation cuvette with a 2-mm gap (NEPAGENE) and transformed by electroporation using a NEPA21 electroporator. Parameters of two pore-forming pulses were set at 300 V with a pulse width, interval, and decay rate of 8 ms, 50 ms, and 10%, respectively, while those of the transfer pulses were set at one polarity-exchanged pulse of 20 V with a pulse width of 50 ms. The transformants were incubated at 25°C for 24 h with gentle shaking in TAP medium containing 40 mM sucrose. After 24-h recovery culture, the cell suspensions were collected, resuspended in fresh TAP medium, and stored under dark conditions until cell sorting experiments. The stored cell suspension was diluted in fresh TAP medium with 10  $\mu\text{M}$  Paromomycin (Wako) and was incubated in culture flasks (working volume: 20 mL) under 14h:10h light:dark cycle with illumination at 120-140  $\mu\text{mol photons/m}^2/\text{s}$  at 25°C for 5 days. After 3-day pre-culture in TAP medium without Paromomycin, the cells were transferred to TAP medium to simulate the high- $\text{CO}_2$  condition where the LCIB-Clover was diffused in the chloroplast. One day later, the sorting experiments were performed with the intelligent IACS. In a separate experiment, we also prepared

cells cultured in Tris-phosphate (TP) medium (Tris-minimal medium, <https://www.chlamycollection.org/methods/media-recipes/tap-and-tris-minimal/>) to simulate the low-CO<sub>2</sub> condition, where the LCIB-Clover was localized around the pyrenoid.

### **Microscopic evaluation of the sorted mutants**

Digital fluorescence and transmission images of BC-9 and sorted mutant cells were acquired with an oil immersion objective lens (HC PL APO 63 × /1.40, Leica) using an inverted laser-scanning confocal fluorescence microscope TCS SP8 (Leica) equipped with a sensitive hybrid detector (HyD). LCIB-Clover was excited at 488 nm and the emission was detected in a wavelength range of 500 - 520 nm. Image scanning was performed with a pinhole size of 0.8 Airy units, at a pixel size of 25 nm, and with a line scan speed of 400 Hz. Contrast adjustment was applied equally to the entire image using LAS X software (Leica).

### **Sorting experiments**

As preparatory steps, we checked and tuned the conditions of the intelligent IACS before sorting experiments. First, to check the hydrodynamic focusing and sorting performance, we visualized the sample stream by flowing highly concentrated 1-μm particles (~1×10<sup>9</sup> particles/mL). While flowing the particles, we activated the sorter with 1,000 pulses/s constantly and recorded movies with the high-speed CMOS camera. The position of the sample stream, the diameter of the sample stream, and the width of the sort window were at the center of the microchannel, 20 μm, and ~400 μs, respectively. Second, to check the stability of the flow in the microchannel, we measured 6-μm particles at the OI1, OI3, and OI4 optical interrogation spots (n > 2,000). The slope of the linear regression line for the OI1-OI3 latency versus the OI1-OI4 latency was used to predict the sort latency. To validate the latency settings, we observed the trajectory of flowing 6-μm sorted particles using the high-speed CMOS camera. After these preparatory steps, we started sorting experiments. The sorting procedure of the intelligent IACS consists of event rate tuning, sorting, and cell/particle collection. After setting the prepared samples to the sample inlet, we injected the sample into the microfluidic chip under a boost mode (sample:sheath1:sheath2 = 3:16:40) for 2 min, followed by a fine focusing mode (1:16:40). Then, we activated the entire intelligent IACS to start sorting. After disposing the initial 1.5-min sample to avoid unwanted carryovers, we started collecting sorted and unsorted cells. The cell-collecting duration depends on the experiment.

### **Evaluation of the purity and yield**

To evaluate the intelligent IACS's sorting yield and purity of sorted cells, we constructed a centrifugation-based cell counting device. A sample sorted with the intelligent IACS was loaded into a custom glass bottom chamber with a volume of 1 mL and a viewing area of 7 mm in diameter fabricated

by bonding a glass substrate (Matsunami, thickness No.1 0.12-0.17 mm, 25 mm diameter) and a cylindrical structure made of PDMS (Dow Corning, polydimethylsiloxane, Sylgard 184), and centrifuged at 300 g for 10 min to collect cells on the glass substrate. To avoid non-specific binding of cells to PDMS, the chamber was incubated with PBS containing with 1% BSA (Bovine serum albumin solution, A9576, sigma) before loading the sample. After the centrifugation, we scanned the whole viewing area with a commercially available fluorescence microscope (Nikon instruments, Ti2) equipped with a digital CMOS camera (Hamamatsu Photonics, ORCA Flash 4.0 V3) and a 20x objective lens (S Plan Fluor ELWD 20x DIC N1) and performed image analysis to count cells with NIS-Elements AR (Nikon instruments). For the polystyrene particle sorting experiment shown in Figures 2F and 2G, 19x19 fields of vision were scanned for each of the collection and waste fluids. The 6- $\mu\text{m}$  and 3- $\mu\text{m}$  particles were detected in the GFP and Cy5 channels and colored in green and magenta, respectively, and then counted with the measurement tool in NIS-elements with a threshold (50-4095 for the GFP channel, 200-4095 for the Cy5 channel) to obtain the number of the sorted and unsorted particles in the collection fluid ( $N_{3-\mu\text{m}, \text{collection}}$  and  $N_{6-\mu\text{m}, \text{collection}}$ ) and in the waste fluid ( $N_{3-\mu\text{m}, \text{waste}}$  and  $N_{6-\mu\text{m}, \text{waste}}$ ), respectively. From the enumeration results, the sort yield and the purity of the sorted particles were calculated as the ratio of  $N_{3-\mu\text{m}, \text{collection}}$  to the sum of  $N_{3-\mu\text{m}, \text{collection}}$  and  $N_{3-\mu\text{m}, \text{waste}}$ , and the ratio of  $N_{3-\mu\text{m}, \text{collection}}$  to the sum of  $N_{3-\mu\text{m}, \text{collection}}$  and  $N_{6-\mu\text{m}, \text{collection}}$ , respectively. For demonstration purposes, we further constructed and optimized large images showing the whole viewing area to visualize individual beads with enhanced contrast and smoothing functions in ImageJ.

## Quantification and Statistical Analysis

### Deep convolutional neural network

The deep CNN used for sorting particles and blood cells is one type of a deep neural network which is a flexible and powerful machine learning method with large datasets (LeCun et al., 2015; LeCun et al., 1989). Since the deep CNN requires a significant amount of computation, it has been considered unsuitable for real-time processing. In order to achieve both high accuracy and real-time operation, we designed a six-layer CNN to classify 3- $\mu\text{m}$  particles from 6- $\mu\text{m}$  particles, which is composed of four convolution layers and two fully connected layers accompanied by two max-pooling and three dropout connections (Figure 2E). We also designed an eight-layer CNN to classify platelet aggregates from single platelets and leukocytes (Figure 5A). The deep CNNs were constructed using TensorFlow (Abadi et al., 2016) and Keras software frameworks and accelerated with the GPU. For training and validating both the deep CNNs, we prepared 8,890 images of 3- $\mu\text{m}$  particles and 9,246 images of 6- $\mu\text{m}$  particles for the six-layer CNN, and 2,106 images of platelet aggregates, 2,073 images of single platelets, and 331 images of leukocytes for the eight-layer CNN, all taken by the FDM microscope and manually classified to construct

class labels as the ground truths, some of which were used for training while the others were used for validation. To validate the reproducibility of the CNN training, we repeated the training with a training dataset in the platelet study four times and achieved an accuracy level of  $> 94\%$  in all the four cases. To implement CNN classifiers for other applications, preparation of the class-labeled grand truth data as well as coding on the TensorFlow and Keras software frameworks are required. We obtained the probability of the image being classified as each of the classes to make a sort decision based on the probabilities. The sum of all the probabilities in each deep CNN is 100%. Figure 5B shows examples of the probabilities in the platelet experiment. We sorted events when the leukocyte-class probability is lower than 1% and the platelet aggregate-class probability is greater than 80% as shown in Figures 5C and 5D indicate the validation results after the training, showing high specificity and sensitivity of 99.0% and 82.0% for the CNN. The trained CNNs were implemented on the IA node. Figures 5E and 5G show the processing time of the image analysis using the deep CNNs, combined with data transfer and the image construction for each event.

### Classical image analysis

For each event, a binary image mask is created from its bright-field image, followed by the extraction of various image features from the bright-field and fluorescence images. Specifically, after applying a median filter to the images to remove noise, Canny edge detection is performed on the bright-field image for detecting the edge of the event (e.g., particle, cell, cell cluster) from the background, and then contours are obtained as a binary image. Based on the extracted contours, the binary image mask of the cell is obtained after morphological operations such as erosion and dilation. The mask is used to extract morphological features of the event including area and shape, as well as intensity information including the average and standard deviation of the bright-field and fluorescence signals. The intelligent IACS can accept any image features such as perimeter, circularity, protein localization, and nucleus-to-cytoplasm ratio as long as the total signal processing time fits within the sort latency. In the *C. reinhardtii* sorting experiment, a CV of the fluorescence signal is calculated as a ratio of the standard deviation to the average of the fluorescence signal intensity. The sort decision is made based on the gating regions shown in Figure 4B.

### Image processing

Processing of the raw images obtained by the FDM microscope is performed in ImageJ. First, 2D mean filtering and 2D minimum filtering are performed for bright-field and fluorescence images, respectively. Then, for the fluorescence images, linear contrast enhancement with histogram stretching is



performed. Finally, a square region around each target cell is cropped for all images. To prepare an overlay image, fluorescence images from FL1 and FL2 are added to a bright-field image.

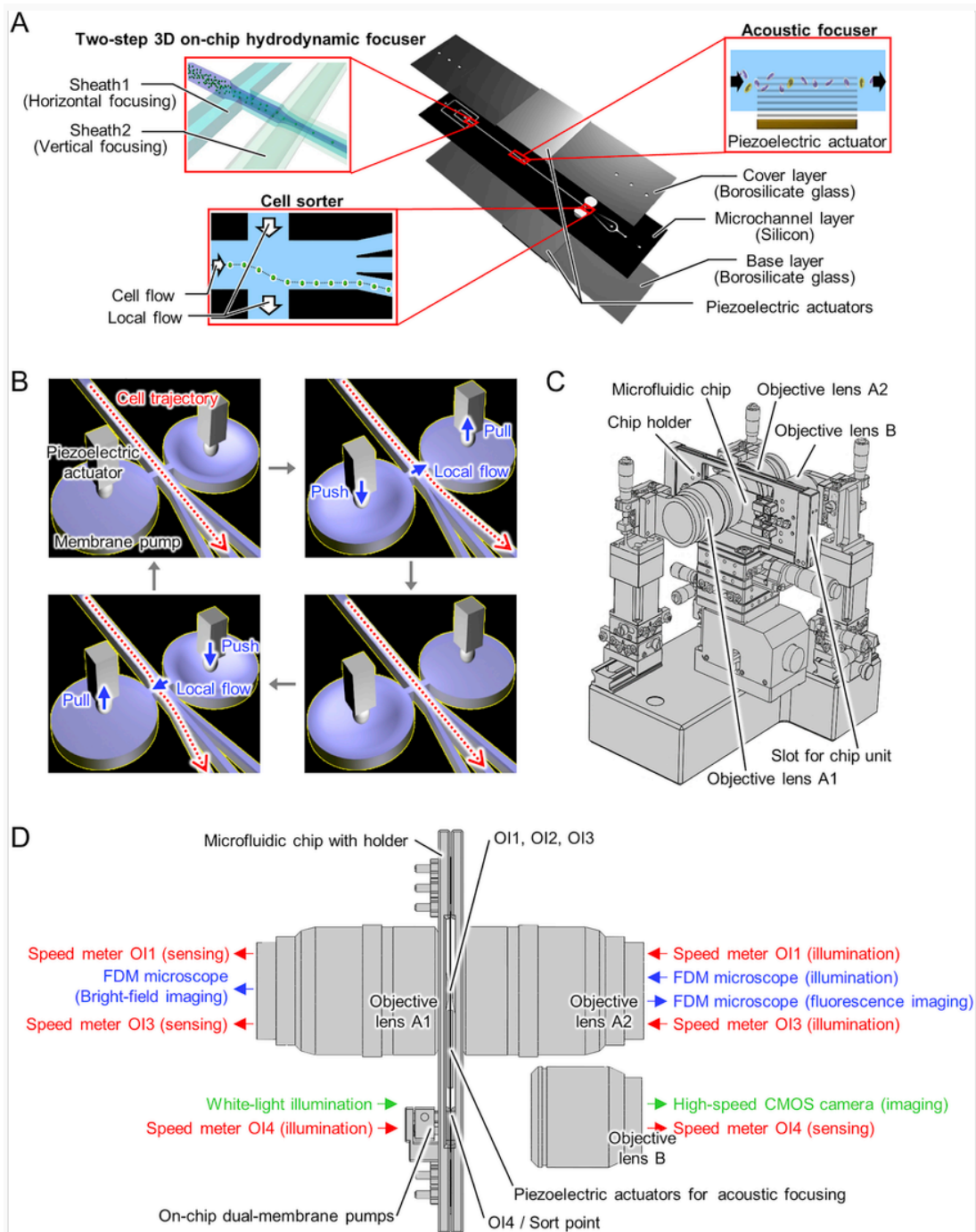
### **Data processing**

The scatterplots, histograms, and processing time plots are drawn using R. The scatterplots are presented as a combination of the contour plots to present a high-density region and dot plots to show the remaining minor events. The processing time plots present events in horizontal axis and processing time in vertical axis, where the measured events are rank-ordered in the total processing time, that is, the event with the minimum total processing time is plotted on the left, followed by all other events in ascending order.

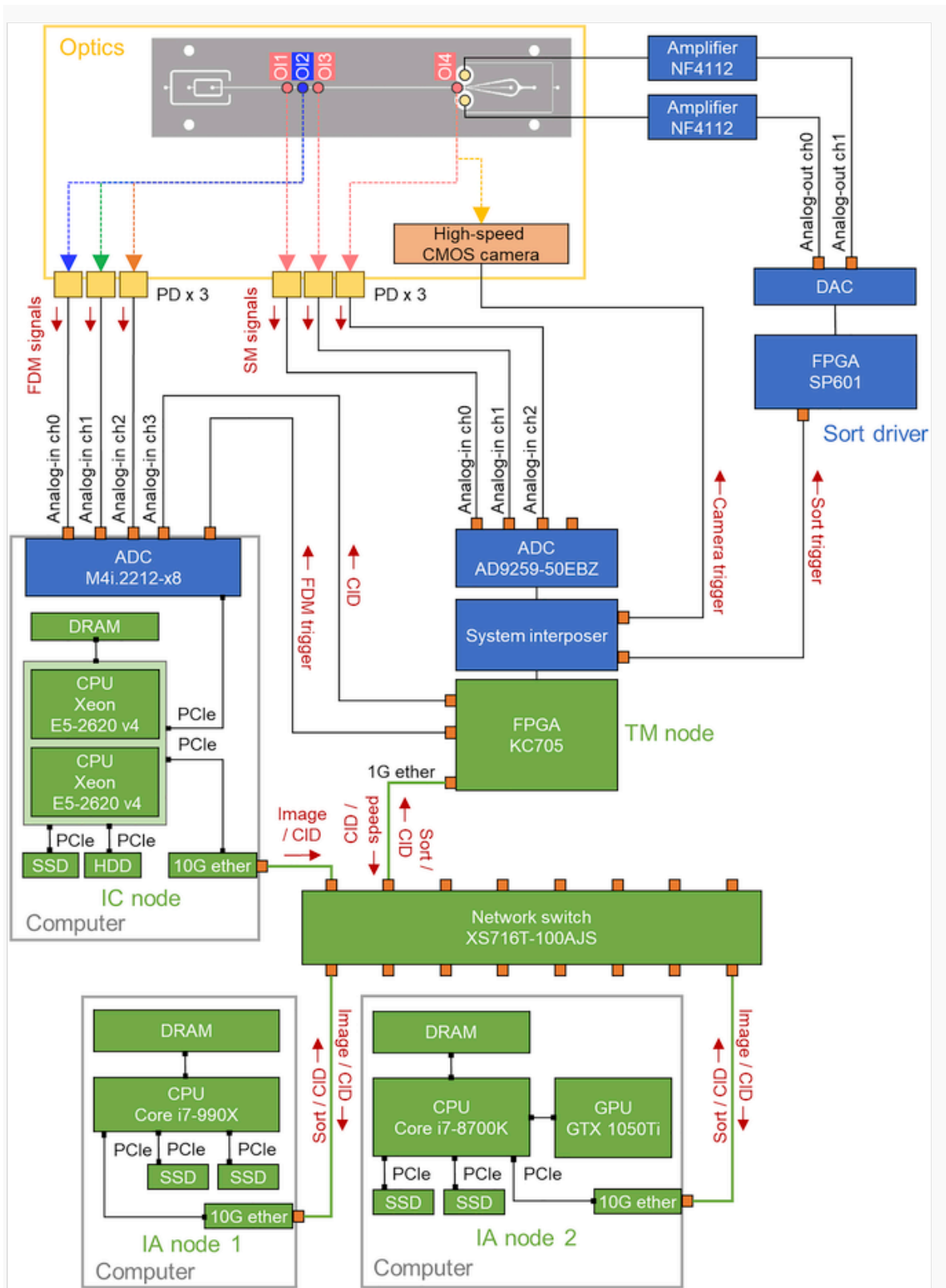
### **Data and Software Availability**

The raw data obtained and the custom codes for image analysis and sort/unsort decision making are available from the corresponding author upon request or from the following webpage: <http://www.goda.chem.s.u-tokyo.ac.jp/intelligentIACS/software.zip>.

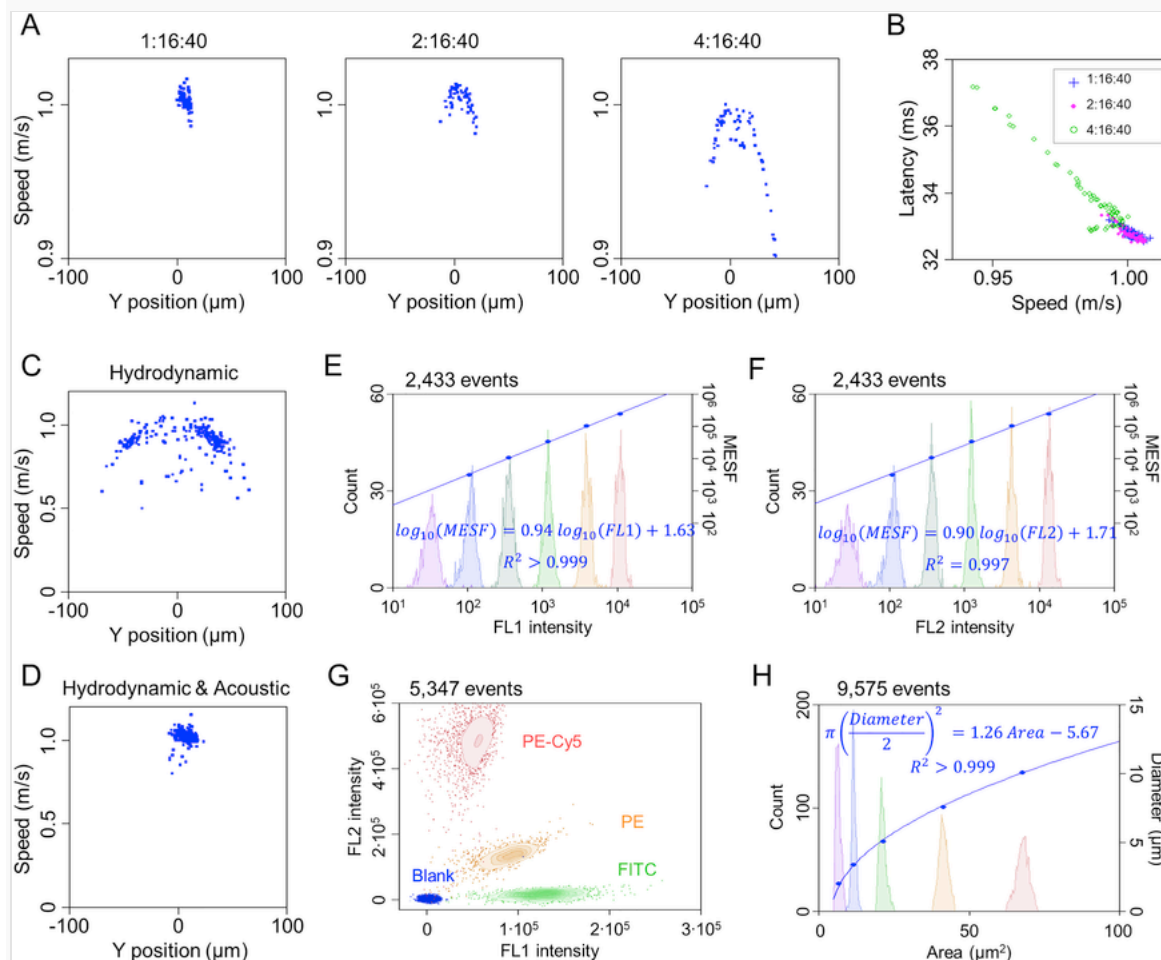
### **Supplementary Figures**



**Figure S1** Details of the Microfluidic Chip and Optics-Microfluidic Integration, Related to Figure 1(A) Schematic of the microfluidic chip.(B) Concept of continuous sorting with the on-chip dual-membrane push-pull sorter.(C) Schematic of the optics-microfluidics integration unit.(D) Installation of the integration unit with details of the optical interrogation (OI1 – OI4) and high-speed CMOS monitoring.

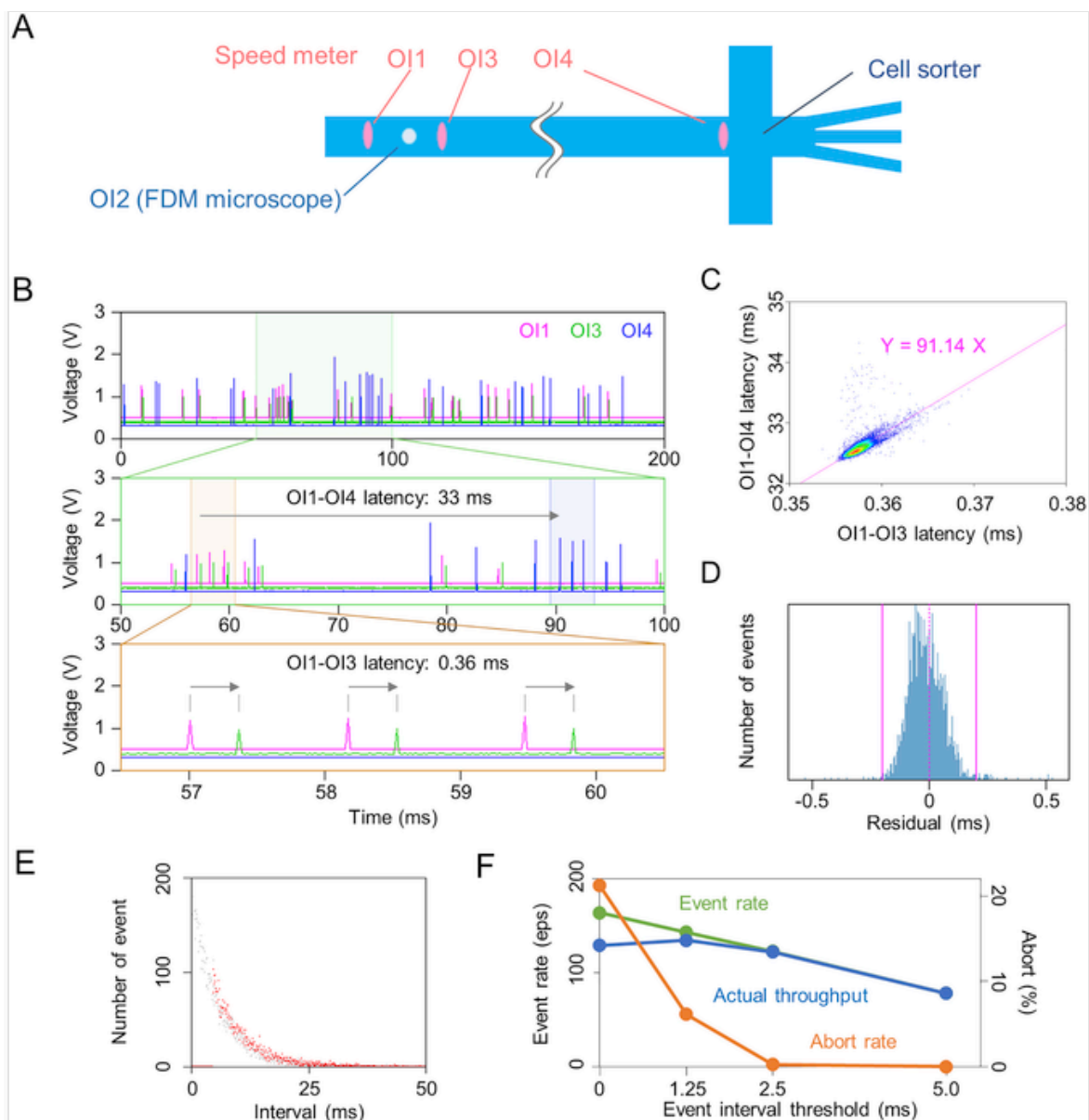


**Figure S2** Real-Time Intelligent Image Processor on the All-IP Network, Related to Figure 1(C): Image construction; IA: Image analysis; TM: Time management; ADC: analog-to-digital converter; DAC: digital-to-analog converter; FPGA: field-programmable gate array; CPU: central processing unit; GPU: graphics processing unit; DRAM: dynamic random-access memory; SSD: solid state drive; HDD: hard disk drive; SM: speed meter; PD: photodetector; CID: cell identifier.



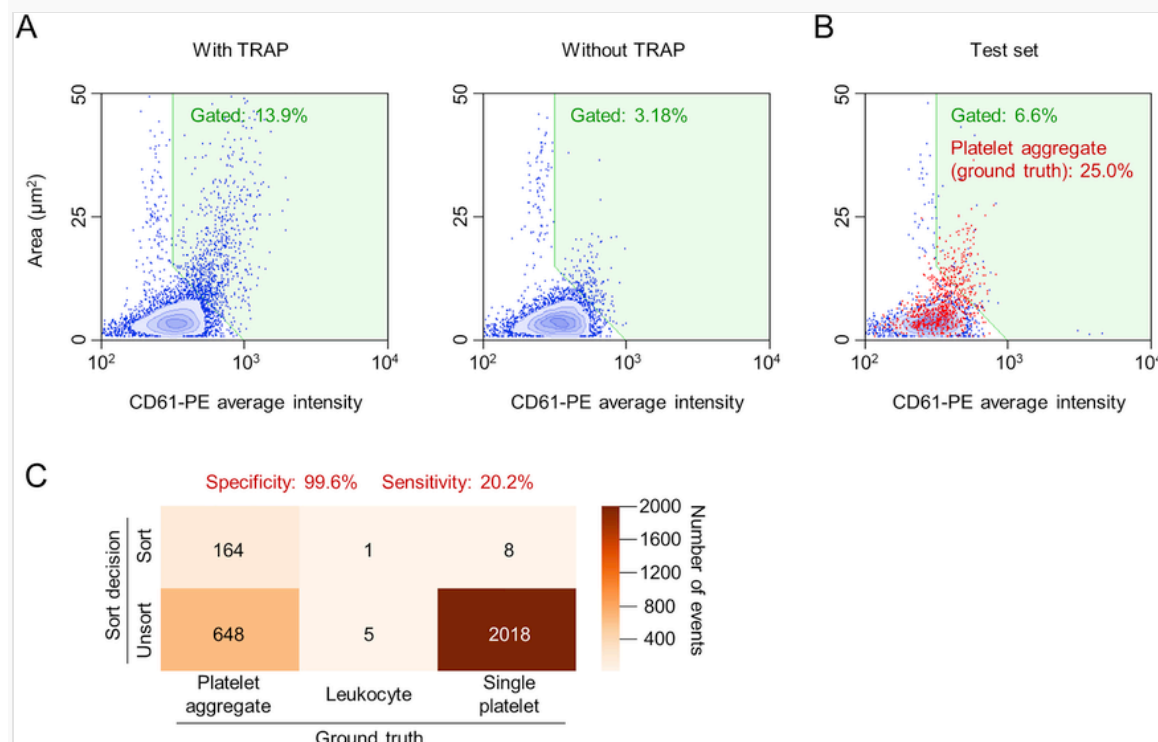
**Figure S3** Characterization of the Microfluidic Chip and FDM Microscope, Related to Figure 1(A) Experimental results of the horizontal position distribution of flowing 6-μm particles (x axis) and the flow speed (y axis) under three different hydrodynamic focusing conditions (sample:sheath1:sheath2). (B) Experimental results of the flow speed and sort latency fluctuations of individual particles under three focusing conditions: sample:sheath1:sheath2 = 1:16:40 (blue), 2:16:40 (magenta), and 4:16:40 (green). (C and D) Experimental results of the horizontal position distribution of flowing *E. gracilis* cells (x axis) and the flow speed under sample:sheath1:sheath2 = 1:16:40, with (D)

and without (C) acoustic focusing.(E and F) Evaluation of the sensitivity, dynamic range, and linearity of the fluorescence detection by FL1 (E) and FL2 (F) using standard 6-peak fluorescent calibration particles. In both panels, each histogram shows a distribution of the signal intensities of the particles with different fluorescence intensities (left axis), whereas the blue dots indicate the average signal intensities and MESF values of the populations (right axis), with a regression line and a fit function shown in blue.(G) Evaluation of the fluorochrome separation using standard compensation particles for FITC, PE, and PE-Cy5, as well as blank particles.(H) Evaluation of the spatial resolution and calibration using size standard particles. The measured size distributions of the 2.0-2.4  $\mu\text{m}$  (purple), 3.0-3.4  $\mu\text{m}$  (blue), 5.0-5.9  $\mu\text{m}$  (green), 7.0-7.9  $\mu\text{m}$  (orange), and 8.0-12.9  $\mu\text{m}$  (red) are shown (left axis) with a regression line and a fit function shown in blue (right axis). In (E), (F), (G), and (H), debris and unfocused images were excluded by applying gating to the area and the bright-field intensity data.



**Figure S4** Characterization of the Speed Meter and Throughput, Related to Figure 1(A) Schematic of the speed meter. OI1 and OI3 are located adjacent to OI2 while OI4 is located slightly upstream of the sorter.(B) Time-course plots of signals from the photodetectors at OI1 (magenta), OI3 (green), and OI4 (blue). Each pulse corresponds to the passage of a particle. The top, middle, and bottom plots show detected particles within a time range of 0 to 200 ms, 50 to 100 ms, and 56.5 to 60.5 ms, respectively. The three sequential peaks highlighted in the middle plot indicate that the latency between OI1 and OI4 is about 33 ms. The OI1 and OI3 peaks are enlarged in the bottom plot, indicating that the latency between OI1 and OI3 is about 0.36 ms.(C) Latency between the OI1 and OI3 points versus the latency

between the OI1 and OI4 points with a linear regression using 6- $\mu$ m particles.(D) Histogram of events in regression residual.(E) Histograms of events in the event-interval time with 169.6 eps with (red) and without (gray) using the digital circuit that aborts events with an interval shorter than a specified duration.(F) Event rate, actual throughput, and abort rate evaluated with the 169.6-eps condition under different event internal threshold values. As the threshold value is increased, the error rate is reduced, whereas the event rate is decreased. Consequently, the actual throughput, which is the product of the event rate and the abort rate, has a peak value at a threshold value of about 1 ms.



**Figure S5** Classical Image Analysis for Detecting Platelet Aggregates, Related to Figure 5(A) Scatterplots of events for blood samples with and without TRAP stimulation in CD61-PE intensity and area. The gating region indicated in green was set to separate TRAP-induced changes.(B) Scatterplot of a test set constructed from multiple measurements combined, with red dots showing platelet aggregates (ground truths).(C) Correlation matrix between the ground truth and the sort decision based on the classical image analysis with the gating region. Due to its poor detection capability, the classical image analysis was not able to identify platelet aggregates with high sensitivity (only 20.2%) for high specificity or vice versa.